

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 8558.13 Seconds  
(without alignments)  
11051.408 Million cell updates/sec

Title: US-09-820-095B-3\_COPY\_1\_2000  
Perfect score: 2000  
Sequence: 1 tctccaaagtccatgggtgcc.....agtcggcggtcagctgcc 2000

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364489745 residues

Total number of hits satisfying chosen parameters: 9053459

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_btgi.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_cv.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2000	100.0	28984	9	AB002059 Homo sapi
2	2000	100.0	147086	9	AC002472 Homo sapi
3	2000	100.0	162470	9	AC007664 Homo sapi
4	1624.6	81.2	175167	2	AC116044 Papio ham
5	1025.4	51.3	147086	9	AC002472 Homo sapi
6	1025.4	51.3	162470	9	AC007664 Homo sapi
7	563	28.1	563	6	CQ073030 Sequence
8	563	28.1	563	6	CQ103845 Sequence
9	563	28.1	563	6	CQ142655 Sequence
10	563	28.1	563	6	CQ178146 Sequence
11	563	28.1	563	6	CQ225890 Sequence
12	563	28.1	563	6	CQ263940 Sequence
13	563	28.1	563	6	CQ301025 Sequence
14	563	28.1	563	6	CQ338301 Sequence
15	554	27.7	554	6	CQ103739 Sequence
16	554	27.7	554	6	CQ178050 Sequence
17	554	27.7	554	6	CQ300915 Sequence
18	458.4	22.9	1904	6	AX714692 Sequence
19	458.4	22.9	1904	9	AK057257 Homo sapi

```

c 20 364.4 18.2 1828 9 BC009991
c 21 288 14.4 288 11 G03785
c 22 158.4 7.9 160 6 AX907126
c 23 158.4 7.9 160 6 BD042659
c 24 141 7.0 792 6 CQ732492
c 25 139 7.0 139 6 CQ116544
c 26 139 7.0 139 6 CQ187685
c 27 139 7.0 139 6 CQ313092
c 28 115.4 5.8 201300 9 AL137073
c 29 114 5.7 3226 9 AK128646
c 30 109.2 5.5 123817 9 AL713980
c 31 109.2 5.5 151964 9 AL663038
c 32 109.2 5.5 174327 2 AL672191
c 33 109.2 5.5 198550 9 AC012170
c 34 106.2 5.4 127967 9 AL353701
c 35 107.4 5.4 72402 9 AL389924
c 36 107.2 5.4 160702 9 AC090937
c 37 107.2 5.4 181321 9 AC093179
c 38 106.6 5.3 136270 9 AC052833
c 39 106.6 5.3 163795 9 AP000356
c 40 106.6 5.3 186858 9 AL356585
c 41 106.4 5.3 141628 9 AL391724
c 42 106.4 5.3 172413 2 AC134388
c 43 106.4 5.3 186044 2 AC084099
c 44 106.4 5.3 191500 2 AC145062
c 45 106.4 5.3 198947 2 AC105308

```

#### ALIGNMENTS

```

RESULT 1
AB002059      28984 bp      DNA      linear      PRI 14-APR-2000
LOCUS      Homo sapiens DNA for Human P2XM, complete cds.
DEFINITION      AB002059.1 GI:2350848
ACCESSION      AB002059.1
VERSION      Human P2XM.
KEYWORDS      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (sites)
AUTHORS      Urano,T., Nishimori,H., Han,H., Furuhata,T., Kimura,Y., Nakamura,Y.
and Tokino,T.
TITLE      Cloning of P2XM, a novel human P2X receptor gene regulated by p53
JOURNAL      Cancer Res. 57 (15), 3281-3287 (1997)
MEDLINE      97384966
PUBMED      9242461
REFERENCE      2 (bases 1 to 28984)
AUTHORS      Nakamura,Y.
TITLE      Direct Submission
JOURNAL      Submitted (22-MAR-1997) Yusuke Nakamura, The Inst. of Medical
Science, The University of Tokyo, Lab. of Molecular Medicine, Human
Genome Center, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
(E-mail:yusuke@ims.u-tokyo.ac.jp, Tel:81-3-5449-5372,
Fax:81-3-5449-5433)
FEATURES
Location/Qualifiers
1..28984
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="22"
/maps="22q11"
join(9106..9239,9843..9993,11889..11960,16575..16650,
16841..16934,17017..17097,17174..17315,17408..17517,
19708..19801,19914..19979,20155..20232,20323..20520)
/gene="HP2XM"
join(9106..9239,9843..9993,11889..11960,16575..16650,
16841..16934,17017..17097,17174..17315,17408..17517,
19708..19801,19914..19979,20155..20232,20323..20520)
/gene="HP2XM"
/codon_start=1

```

```

/product="Human P2XM"
/protein_id="BAA22047.1"
/db_xref="GI:2350849"
/translation="MGSPCATGCGGLDYKTEKVMYTRNRYGALQRLILQFIVVYV
GALLAKGYQREDLEPQPSLIITKLKGVSTQIKELGNRLMDVDFVPPQGENVFL
VTNFLTPAQVQRCRCEHPSVPLANCWVDEDCPEGEGGTHSHGVTGCVFNGFHT
CEIWSWCPVSGVSPRLAQAQNTLPIKNTVTFKFNFSKSNALBTWPTFPKH
RVEPQSPYCPVFRIGDLVAKAGTFFDLALIGSVGIRVHWDCLDITDGDSCWPHYS
FOLEKSYNFRATHWEPQGVBEARTLLKYGIRFDILVTGQAGKFGILPTAVTLGTG
AWLGVVTFECLLILYVDREAHFYWRTKVEBAKAPKATANSWRELALASQARLAE
LRSSAPAPATAAGSQQTCPGWPSPSSDTHLPTHSGSL"
exon
9106..9239
/gene="HP2XM"
/number=1
9843..9993
/gene="HP2XM"
/number=2
11889..11960
/gene="HP2XM"
/number=3
16575..16650
/gene="HP2XM"
/number=4
16841..16934
/gene="HP2XM"
/number=5
17017..17097
/gene="HP2XM"
/number=6
17174..17315
/gene="HP2XM"
/number=7
17408..17517
/gene="HP2XM"
/number=8
19708..19801
/gene="HP2XM"
/number=9
19914..19979
/gene="HP2XM"
/number=10
20155..20232
/gene="HP2XM"
/number=11
20323..20520
/gene="HP2XM"
/number=12
ORIGIN
Query Match      100.0%; Score 2000; DB 9; Length 28984;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1  TCTCAAGTCATGGTCCCTGGTAGGACAGAGGAGATGATGAAACCCCTGCATGGC 60
DB      7067 TCTCAAGTCATGGTCCCTGGTAGGACAGAGGAGATGATGAAACCCCTGCATGGC 7126
QY      61  TATAGCCACCTCCCTCCCTCCCTCCCTGCATCACTACCTGGCCCTATTTTGGCTCTAG 120
DB      7127 TATAGCCACCTCCCTCCCTCCCTCCCTGCATCACTACCTGGCCCTATTTTGGCTCTAG 7186
QY      121 AAGCACTCTCTATGCTCTCTAGGACCACTGCGCGCATATGACAGATAAGAACATCGA 180
DB      7187 AAGCACTCTCTATGCTCTCTAGGACCACTGCGCGCATATGACAGATAAGAACATCGA 7246
QY      181 GGCTAAGCAACGCAATCTTTTCCCTTAAAGTCATACAGCTGTCAAAGAAAGCTGACA 240
DB      7247 GGCTAAGCAACGCAATCTTTTCCCTTAAAGTCATACAGCTGTCAAAGAAAGCTGACA 7306
QY      241 ACCTGGGCAACATAGCGAGATAAAAAATTTTAAATTAGCCAGATGTGGTAGCCCCCTG 300
DB      7307 ACCTGGGCAACATAGCGAGATAAAAAATTTTAAATTAGCCAGATGTGGTAGCCCCCTG 7366

```

```

QY      301 TAGTCTCAGCGACTCAGAGGCTGAGSCAGAGGCTCACCCAGAGTCAGAGTTCAAGGAT 360
DB      7367 TAGTCTCAGCGACTCAGAGGCTGAGSCAGAGGCTCACCCAGAGTCAGAGTTCAAGGAT 7426
QY      361 GCAGTGAAGTATGATCTCTGACCTGACCTGACCTGAAAGCTGGGTGACAGAGCAAGACCTTGGCT 420
DB      7427 GCAGTGAAGTATGATCTCTGACCTGACCTGAAAGCTGGGTGACAGAGCAAGACCTTGGCT 7486
QY      421 CTAATAAATGAATACATAAAGTCTCAGAGTGTGAGTGTAGTGTATCTTCCAGAGTCAGGC 480
DB      7487 CTAATAAATGAATACATAAAGTCTCAGAGTGTGAGTGTAGTGTATCTTCCAGAGTCAGGC 7546
QY      481 CTCTACCTGTCTGATGACAAATGGCACATATATCTCTTTTAACTGATTCAGACACACAAA 540
DB      7547 CTCTACCTGTCTGATGACAAATGGCACATATATCTCTTTTAACTGATTCAGACACACAAA 7606
QY      541 TGTTTTGTGAATATTTTCCCCAGGGAATAAACCGGAAGTAGTTCTAAATCTTATATATCC 600
DB      7607 TGTTTTGTGAATATTTTCCCCAGGGAATAAACCGGAAGTAGTTCTAAATCTTATATATCC 7666
QY      601 ATTATATTAGTTTACCTGTGGAATGGGAAACCCAGCTCTGATTCATTCAGGCGCGG 660
DB      7667 ATTATATTAGTTTACCTGTGGAATGGGAAACCCAGCTCTGATTCATTCAGGCGCGG 7726
QY      661 ACAGCCTTTGGTGACACTGTCTGGCGGGAATTTTCATTTTAACTCTCTTAGAAGCGCT 720
DB      7727 ACAGCCTTTGGTGACACTGTCTGGCGGGAATTTTCATTTTAACTCTCTTAGAAGCGCT 7786
QY      721 TCTCATGTGTAAGTTCCTGATGCGGACGAGCGGCGAGAGAGGAGGCGGCGGCTGGAGA 780
DB      7787 TCTCATGTGTAAGTTCCTGATGCGGACGAGCGGCGAGAGGAGGAGGCGGCGGCTGGAGA 7846
QY      781 CGCCCCGACAGAGGCTACGTGCCCTGTGTCGACAGAGGTCTCTCTGCTCTCTCGCGCGCGCC 840
DB      7847 CGCCCCGACAGAGGCTACGTGCCCTGTGTCGACAGAGGTCTCTCTGCTCTCTCGCGCGCGCC 7906
QY      841 AGCCACACTCCACACACCCCTGGGGGAGAGAGCCCAAGGGGAGAGAGCGGCGCTGGGCC 900
DB      7907 AGCCACACTCCACACACCCCTGGGGGAGAGAGCCCAAGGGGAGAGAGCGGCGCTGGGCC 7966
QY      901 CTGCCCCGACAGACCTTTCCCTCTCTAGGTGCGAGTCTGAATCGGCTTGGGACCTGCTTG 960
DB      7967 CTGCCCCGACAGACCTTTCCCTCTCTAGGTGCGAGTCTGAATCGGCTTGGGACCTGCTTG 8026
QY      961 GCTTGGGGACCCCTGCAAGACGTCCAGAGCGCGCGCTGCGCTCTCTCTCTGCTTTTATA 1020
DB      8027 GCTTGGGGACCCCTGCAAGACGTCCAGAGCGCGCGCTGCGCTCTCTCTCTGCTTTTATA 8086
QY      1021 TCCTCCCGACAGACTCTGGCAGGAACCGCTCATCGTTACGCCCTTTTCGACGCTCAGACC 1080
DB      8087 TCCTCCCGACAGACTCTGGCAGGAACCGCTCATCGTTACGCCCTTTTCGACGCTCAGACC 8146
QY      1081 CTGAGCGGAGACCCGCTTGGCGCTCACTTAGAGCGCAACCCGGGATGTGGGGGAGTCTC 1140
DB      8147 CTGAGCGGAGACCCGCTTGGCGCTCACTTAGAGCGCAACCCGGGATGTGGGGGAGTCTC 8206
QY      1141 TCGCGCTGCGCTGACCAATCGAGTGTGGCTCCATCGACTGGCTCTGCCACGCAATTA 1200
DB      8207 TCGCGCTGCGCTGACCAATCGAGTGTGGCTCCATCGACTGGCTCTGCCACGCAATTA 8266
QY      1201 GCGACGCGCTCCCGCGCGCGGTGCGCCCGCAACCCAGTGTCTGAGTGTGCCGTAGAAA 1260
DB      8267 GCGACGCGCTCCCGCGCGCGGTGCGCCCGCAACCCAGTGTCTGAGTGTGCCGTAGAAA 8326
QY      1261 CGGTGGCTCTCTGCGCTGAGGCTCTCTGCGCTGAGAGGATAAATGACACGCGCACCGGC 1320
DB      8327 CGGTGGCTCTCTGCGCTGAGGCTCTCTGCGCTGAGAGGATAAATGACACGCGCACCGGC 8386
QY      1321 TATGACCTGGGCTGGCGCGCTTGTGGGCATCTCTCTGCTCTCTAGGGGCTTCCAGCAT 1380
DB      8387 TATGACCTGGGCTGGCGCGCTTGTGGGCATCTCTCTGCTCTCTAGGGGCTTCCAGCAT 8446
QY      1381 CGCCCCCTTTCTGGGACTGGGAAACAGCGCTGACTCCAGGACTTGTGTCTCTACTG 1440

```

Db 8447 CGCCCCCTTTCTGTGACTGGGAAACACGCTGACTCCAGGACTTGTGTCTCTCAGCTG 8506  
QY 1441 CACTGGGGAGGTGGCGGGGCGCAGCTTTTTCAGAGAGGCGCTGGGGAATCTTCGAGAGCCAG 1500  
Db 8507 CACTGGGGAAGGTGGCGGGGCGCAGCTTTTTCAGAGAGGCGCTGGGGAATCTTCGAGAGCCAG 8566  
QY 1501 GTCAACCTCTCATCTGTGCTCTTAGTTATCTTGTATCTTGTGATCTCTGCTTTGTATAGCTG 1560  
Db 8567 GTCAACCTCTCATCTGTGCTCTTAGTTATCTTGTATCTTGTGATCTCTGCTTTGTATAGCTG 8626  
QY 1561 CTCCTCTGCACAGGAACCTCCATCCCATCTTGTCTGTCTGTGCAACTTCAGAAATCTG 1620  
Db 8627 CTCCTCTGCACAGGAACCTCCATCCCATCTTGTCTGTCTGTGCAACTTCAGAAATCTG 8686  
QY 1621 CAAGGTCAGCTTAGAGTCATCTTTCGGGAAGCTTTCTCAACCTCCCGCCCTG 1680  
Db 8687 CAAGGTCAGCTTAGAGTCATCTTTCGGGAAGCTTTCTCAACCTCCCGCCCTG 8746  
QY 1681 CTCTGCTGCGCTCAGGCGCTCTCTCACAGCACTGATACAGCTCTCGTCTCCACCT 1740  
Db 8747 CTCTGCTGCGCTCAGGCGCTCTCTCACAGCACTGATACAGCTCTCGTCTCCACCT 8806  
QY 1741 CCCACCACTCTCACTCCACCCCGAGGAAGTGAGGCCAGAGGCGAGGACAGAGCTCTCTC 1800  
Db 8807 CCCACCACTCTCACTCCACCCCGAGGAAGTGAGGCCAGAGGCGAGGACAGAGCTCTCTC 8866  
QY 1801 TGTCTCTGTCTGCCAGGCGCCAGCAAGGATGTAGGAGGTGGAGGTGCAGGGCA 1860  
Db 8867 TGTCTCTGTCTGCCAGGCGCCAGCAAGGATGTAGGAGGTGGAGGTGCAGGGCA 8926  
QY 1861 GCTGGGATTAGGGGTTGAGGCGCTGGGTGTGGAGGCTGGATCTGGATCTCTTTAGTGG 1920  
Db 8927 GCTGGGATTAGGGGTTGAGGCGCTGGGTGTGGAGGCTGGATCTGGATCTCTTTAGTGG 8986  
QY 1921 AAGTGCTCTTTAAGCAAGCACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1980  
Db 8987 AAGTGCTCTTTAAGCAAGCACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 9046  
QY 1981 AGCTGGCGCTGCAGCTGCCA 2000  
Db 9047 AGCTGGCGCTGCAGCTGCCA 9066

RESULT 2  
AC002472 147086 bp DNA linear PRI 09-APR-2003  
LOCUS Homo sapiens Chromosome 22q11.2 PAC Clone p\_n5 In BCRL2-GGT Region,  
complete sequence.  
AC002472  
VERSION AC002472.8 GI:24137490  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 147086)  
AUTHORS Budarf,M.L. and Emanuel,B.S.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Lao,V. and Roe,B.A.  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 147086)  
AUTHORS Zhang,G., Lao,V., Zhan,M. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (20-AUG-1997) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 4 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (19-SEP-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA

REFERENCE 5 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (23-OCT-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 6 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (13-NOV-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 7 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (18-NOV-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 8 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (21-NOV-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 9 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (14-APR-1999) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 10 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (13-SEP-1999) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 11 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (13-MAY-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 12 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 13 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (20-MAY-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 14 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (31-MAY-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 15 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (01-SEP-2002) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 16 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (19-OCT-2002) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA

REFERENCE	OK 73019, USA	Db	58035	TCTCATGTAAAGTTCTGATGCCGACGAGCGCCGAGAGAGGCGAGGGCTGGAGA	58094
AUTHORS	17 (bases 1 to 147086)	Qy	781	CGCCCCGACAGAGGCTACGTGCCCTCTGACAGAGGTCTCTGCTCTCTCGCGCGCGCC	840
TITLE	Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.	Db	58095	CGCCCCGACAGAGGCTACGTGCCCTCTGACAGAGGTCTCTGCTCTCTCGCGCGCGCC	58154
JOURNAL	Submitted (09-APR-2003) Department of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy	841	AGCCCACTCCCAACACCCCTGCGGAGAACCCCAAGGGAGGAGAGCGGCTGGCCC	900
COMMENT	On Oct 19, 2002 this sequence version replaced gi:22597497. Because these overlapping clones came from different libraries.	Db	58155	AGCCCACTCCCAACACCCCTGCGGAGAACCCCAAGGGAGGAGAGCGGCTGGCCC	58214
FEATURES	Location/Qualifiers	Qy	901	CTGCCCGAGACACTTCCGCTCTTAGGTGGAAGTCTGAATCGGCTTGGGACCTGCTTG	960
source	1. .147086 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosomes="22" /map="22q11.2" /clone="p_n5"	Db	58215	CTGCCCGAGACACTTCCGCTCTTAGGTGGAAGTCTGAATCGGCTTGGGACCTGCTTG	58274
ORIGIN		Qy	961	GCTTGGGGACCCCTGCAAGACGTCACAGGCGCGCTTCCCTCTCTCTCTCTCTCTCT	1020
Query Match	100.0%; Score 2000; DB 9; Length 147086;	Db	58275	GCTTGGGGACCCCTGCAAGACGTCACAGGCGCGCTTCCCTCTCTCTCTCTCTCTCT	58334
Best Local Similarity	100.0%; Pred. No. 0;	Qy	1021	TCTTCCCGACACTCTGGGAGAACCGCTCATGTTACGCCCCCTTTTCGAGCCTCAGACC	1080
Mismatches 2000; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	Db	58335	TCTTCCCGACACTCTGGGAGAACCGCTCATGTTACGCCCCCTTTTCGAGCCTCAGACC	58394
1	TCTCCAAATCCATGGTGGCTGGTAGGACAGAGGGGGATGATGTGAACCCCTGCATGCC	Qy	1081	CTGAGCGGAGACCGCTTGGCGCTCCTACTTAGAGCGGACCGCGGATGTGGCGGAGTC	1140
57315	TCTCCAAATCCATGGTGGCTGGTAGGACAGAGGGGGATGATGTGAACCCCTGCATGCC	Db	58395	CTGAGCGGAGACCGCTTGGCGCTCCTACTTAGAGCGGACCGCGGATGTGGCGGAGTC	58454
61	TATAGCCACTGCTCTCCCTGCCCTGCATCACTACTGCGCTATTTTGTGCTCTAG	Qy	1141	TGCGGCTGCGCTGACCAATCGAGTGTGGCTCCATCGACTGGCGTCTGCCAGCGCAATTA	1200
57375	TATAGCCACTGCTCTCCCTGCCCTGCATCACTACTGCGCTATTTTGTGCTCTAG	Db	58455	TGCGGCTGCGCTGACCAATCGAGTGTGGCTCCATCGACTGGCGTCTGCCAGCGCAATTA	58514
121	AAGCACTGCTCTATGCTCTTAGACCACTGCGCGCATATGACAGATAAGACATCGA	Qy	1201	GCCAGCGGCTCCCGCGGCGGTGCGCCCGCAACCCAGTGTGTAGTGTGCGGTAGAAA	1260
57435	AAGCACTGCTCTATGCTCTTAGACCACTGCGCGCATATGACAGATAAGACATCGA	Db	58515	GCCAGCGGCTCCCGCGGCGGTGCGCCCGCAACCCAGTGTGTAGTGTGCGGTAGAAA	58574
181	GGCTAAGGCAACGCAATCTTTTCTTAAAGTCATACAGCTGTCTAAAGAAAGCTGGACA	Qy	1261	CCCTGCTCTCTCTGCGCTGAGGCTCTCGCCTGAGAGGATAAATCGACGCGCCACGGGC	1320
57495	GGCTAAGGCAACGCAATCTTTTCTTAAAGTCATACAGCTGTCTAAAGAAAGCTGGACA	Db	58575	CCCTGCTCTCTCTGCGCTGAGGCTCTCGCCTGAGAGGATAAATCGACGCGCCACGGGC	58634
241	ACCTGGCAACATAGCAGATAAAAATTTATTTAAATTTAGCAGATGTGGTACCCCTG	Qy	1321	TATGCACTGGCTGGCGCTTGTGGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1380
57555	ACCTGGCAACATAGCAGATAAAAATTTATTTAAATTTAGCAGATGTGGTACCCCTG	Db	58635	TATGCACTGGCTGGCGCTTGTGGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	58694
301	TAGTCTCAGGCTCAGGAGGTGAGGAGGAGGCTCACAGAGTGCAGAGTTCAGAGAT	Qy	1381	CGCCCCCTTTCTGCTGACTGGGAAACACGCTTACCTCCAGGACTTGTGTGTCTCTCTG	1440
57615	TAGTCTCAGGCTCAGGAGGTGAGGAGGAGGCTCACAGAGTGCAGAGTTCAGAGAT	Db	58695	CGCCCCCTTTCTGCTGACTGGGAAACACGCTTACCTCCAGGACTTGTGTGTCTCTCTG	58754
361	GCAGTGAGCTATGATCCTGCCACTGCATGAAAGCTGGGTGACAGAGCAAGACCTGGCT	Qy	1441	CACTGGGAGAGTGGCGGGGAGCTTTTTCAGAGGCGCTGGGAACTTCGAGAGCCAG	1500
57675	GCAGTGAGCTATGATCCTGCCACTGCATGAAAGCTGGGTGACAGAGCAAGACCTGGCT	Db	58755	CACTGGGAGAGTGGCGGGGAGCTTTTTCAGAGGCGCTGGGAACTTCGAGAGCCAG	58814
421	CTAATAATGATATACATAAAGTCTCAGAGTGTGGTGTAGTCTAATCTCCAGAGTCAGC	Qy	1501	GTCAACCTCTCATCTGCTGCTTGTAGTTATCTTGATGCTCTGCTGCTGCTGCTGCTG	1560
57735	CTAATAATGATATACATAAAGTCTCAGAGTGTGGTGTAGTCTAATCTCCAGAGTCAGC	Db	58815	GTCAACCTCTCATCTGCTGCTTGTAGTTATCTTGATGCTCTGCTGCTGCTGCTGCTG	58874
481	CTCTACCTGTCTGATGACAAATGGACACTATGCTTTTAAACCTGATTGACAGACCAAA	Qy	1561	CTCCCTGCACACGAAACCTCCATCCCATCTTTGTCTGCTGCTGCTGCTGCTGCTGCTG	1620
57795	CTCTACCTGTCTGATGACAAATGGACACTATGCTTTTAAACCTGATTGACAGACCAAA	Db	58875	CTCCCTGCACACGAAACCTCCATCCCATCTTTGTCTGCTGCTGCTGCTGCTGCTGCTG	58934
541	TGTTTTGTGAATATTTTCCAGAGGAAACCCGGAAGTAGTTCTAATTTCTATACATCC	Qy	1621	CAAGGTGAGCTTAGAGTCACTTCTTCGGAAGCTTCTCTCAACACCTCCCGGCTG	1680
57855	TGTTTTGTGAATATTTTCCAGAGGAAACCCGGAAGTAGTTCTAATTTCTATACATCC	Db	58935	CAAGGTGAGCTTAGAGTCACTTCTTCGGAAGCTTCTCTCAACACCTCCCGGCTG	58994
601	ATTATATTAGTTTACCTGTGATTTGGGAAACCCGGAAGTAGTTCTAATTTCTATACATCC	Qy	1681	CTGCTGCTCCCTCAGGCTCTCTCTCAGACACTGATAACAGCTCTCGCTCCACCTCT	1740
57915	ATTATATTAGTTTACCTGTGATTTGGGAAACCCGGAAGTAGTTCTAATTTCTATACATCC	Db	58995	CTGCTGCTCCCTCAGGCTCTCTCTCAGACACTGATAACAGCTCTCGCTCCACCTCT	59054
661	ACAGCCTTTGGTGCATCTGTCTGGCGGATTTTCAATTTTAACTCTCTTAGAGAGCGCT	Qy	1741	CCACACCTCTCATCTCCGACCGAGAGTGAAGGCGAGGCGCAGAGCTGTCTGTGC	1800
57975	ACAGCCTTTGGTGCATCTGTCTGGCGGATTTTCAATTTTAACTCTCTTAGAGAGCGCT	Db	59055	CCACACCTCTCATCTCCGACCGAGAGTGAAGGCGCAGAGCTGTCTGTGTGC	59114
721	TCTCATGTGTAAGTTCTGTGCGCCAGGAGCGCCGAGGAGGCGAGGGGCTGGAGA	Qy	1801	TGTTCTCTGTGTGCCAGGCGCCGACAAAGGGAAATGTAGGGAGGTGGGAGGTGCAGGCA	1860
		Db	59115	TGTTCTCTGTGTGCCAGGCGCCGACAAAGGGAAATGTAGGGAGGTGGGAGGTGCAGGCA	59174







-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 11 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\*  
\* 1 6404: contig of 6404 bp in length  
\* 6405 6504: gap of 100 bp  
\* 6505 12196: contig of 5692 bp in length  
\* 12197 12396: gap of 100 bp  
\* 12397 15159: contig of 2863 bp in length  
\* 15160 15259: gap of 100 bp  
\* 15260 25315: contig of 10056 bp in length  
\* 25316 25415: gap of 100 bp  
\* 25416 79413: contig of 5398 bp in length  
\* 79414 79513: gap of 100 bp  
\* 79514 89805: contig of 10292 bp in length  
\* 89806 89905: gap of 100 bp  
\* 89906 92601: contig of 2696 bp in length  
\* 92602 92701: gap of 100 bp  
\* 92702 101172: contig of 8471 bp in length  
\* 101173 101272: gap of 100 bp  
\* 101273 114938: contig of 13666 bp in length  
\* 114939 115039: gap of 100 bp  
\* 115039 141113: contig of 26075 bp in length  
\* 141114 141214: gap of 100 bp  
\* 141214 175167: contig of 33954 bp in length.  
\*  
\* Location/Qualifiers  
\* 1. 175167  
\* /organism="Papio hamadryas"  
\* /mol\_type="genomic DNA"  
\* /db\_xref="taxon:9557"  
\* /clone="RP41-7017"  
\* /clone\_lib="RPCI-41 Male Baboon BAC"

FEATURES

source

ORIGIN

Query Match 81.2%; Score 1624.6; DB 2; Length 175167;  
Best Local Similarity 91.2%; Pred. No. 0;  
Matches 1852; Conservative 0; Mismatches 129; Indels 50; Gaps 10;  
  
QY 1 TCTCCAAAGTCCATGGGTGCTGTAGGAGACAGGGGGATGAATGTGAACCCCTGATGGC 60  
DB 143455 TCTCCAAAGCCTAGGGTGCCTGTAGGAGACAGGGGGATGAATGTGAACCCCTGATGGC 143514  
  
QY 61 TATAGCCACTGCTGCTCCCTCCCTGCTGATCACTACCTGGCTATTTTGGCTCTAG 120  
DB 143515 TATAGCCACTGCTGCTCCCTCCCTGCTGATCACTACCTGGCTATTTGCTCTAG 143574  
  
QY 121 AAGCACTGCTTCTATGCTCTCTTAGGACCACTCCCGCATATGACAGATAAGAAATCGA 180  
DB 143575 AAGCACTGCTTCTATGCTCTCTTAGGACCACTCCCGCATATGACAGATAAGAAATCGA 143634  
  
QY 181 GGCTAAGGCAACGCAATCTTTTCCCTTAAAGTCATACAGCTGTCAAAGAAAGCTGGACA 240  
DB 143635 GGCTAAGGCAACGCAATCTTTTCCCTTAAAGTCATACAGCTGTCAAAGAAAGCTGGACA 143694  
  
QY 241 ACCTGGGCAACATAGCGAGATAAATAATTTTAAATATAGCCAGATGTGTAGCCCTCG 300  
DB 143695 ACCTGGGCAACATAGCGAGATAAATAATTTTAAATATAGCCAGATGTGTAGCCCTCG 143754  
  
QY 301 TAGTCTCAGGCACTCAGAGGCTGAGGACAGGAGGCTCACCAGAGTCAGAGTTCAGAGAT 360  
DB 143755 TAGTCTCAGGCACTCAGAGGCTGAGGACAGGAGGCTCACCAGAGTCAGAGTTCAGAGAT 143814  
  
QY 361 GCAGTGAAGCTAT-GATCCTGCACCTGCACTGAAGCTGGGTGACAGCAAGACCTGGC 419  
DB 143815 GCAGTGAAGCTATGGATCTCCCACTGCACTGAATCTGGGTGACAGCAAGACCTGGC 143874  
  
QY 420 TCTAATAAATGA-----ATACATAAAGTCTCACAGCTAGTGGTAGCTAATCTCCAGAGT 475

Db 143875 TCTAATAAATGAATTCATAAATAAAGTACCACAGCTAGTGGTAGCAATCTCCAGAGT 143934  
QY 476 CAGGCTCTACCTGTCTGATGACAAATGGCAGACTATGCTTTTAACTGATTCGAGACC 535  
Db 143935 CAGGCTCTACCTGTCTGATGACAAATGGCAGACTATGCTTTTAACTGATTCGAGACC 143994  
QY 536 ACAATGCTTTGTGAATATTTCCAGGAGAAAAACCGAGAGTAGTTCCTAATTTCTATA 595  
Db 143995 ACAATGCTTTGTGAATATTTCCAGGAGAAAAACCGAGAGTAGTTCCTAATTTCTATA 144052  
QY 596 CATCATATATATAGTTTACCTGTGGATTGGGAAAAACCGAGAGTAGTTCCTAATTTCTATA 655  
Db 144053 CATCATATAGTTTACCTGTGGATTGGGAAAAACCGAGAGTAGTTCCTAATTTCTATA 144112  
QY 656 GCGGACAGCCTTTGGTGCACTGTCTGGCGGGATTTTCCATTTTAACTCTCTCTAGAAG 715  
Db 144113 CCGGAGAGCCCTTTGGTGCACTGTCTGGCGGGATTTTCCATTTTAACTCTCTCTAGAAG 144172  
QY 716 GCGCTTCTCATGTAAAGTTTCTGATCCCGCAGGAGCGCGAGAGAGGAGGAGGAGGCT 775  
Db 144173 GCGCTTCTCATGTAAAGTTTCTGATCCCGCAGGAGCGCGAGAGAGGAGGAGGAGGCT 144232  
QY 776 GGAGACGCGCGGAGAGGAGGCTACGTGCGCTGTGGACAGAGGCTCTCTGCTCTCTCGCG 835  
Db 144233 GCAGACGCGCGGAGAGGAGGCTACGTGCGCTGTGGACAGAGGCTCTCTGCTCTCTCGCG 144289  
QY 836 GCGCAGCAGCCTCCCAACACCCCTCGGGAGAGAGCCCAAGGGGAGGAGAGGAGGCT 895  
Db 144290 GCGCAGCAGCCTCCCAACACCCCTCGGGAGAGAGCCCAAGGGGAGGAGAGGAGGCT 144349  
QY 896 GCGCCTCTGCGCGGAGCAGCCTCTGCTCTCTGATGCGAGTCTGATGCGCCTTGGAGCCT 955  
Db 144350 GCGCCTCTGCGCGGAGCAGCCTCTGCTCTCTGATGCGAGTCTGATGCGCCTTGGAGCCT 144401  
QY 956 GCTTGGCTTGGGAGACCCCTGCAAGAGCTCCACAGGCGCGCTGCTCTCTCTCTCTGCT 1015  
Db 144402 GCTTGGCTTGGGAGACCCCTGCAAGAGCTCCACAGGCGCGCTGCTCTCTCTCTCTGCT 144461  
QY 1016 TTTATCTCTCCAGAGCCTCTGGCAGGAAACCGCTCATCTGTTAC----- 1058  
Db 144462 CTTTATCTCTCCAGAGCCTCTGCGAGGAAACCATTTCTGTTGTTTAAAGCCCTTTTCTCAGGTTA 144521  
QY 1059 -GCCCTTTGCGAGCCTCAGACCTGAGGCGGAGACCGCTTGGCGCTCTCCTTAGAGCGC 1117  
Db 144522 AGCCCTTTGCGAGCCTCAGACCTCAGGCGGAGACCGCTTGGCGCTCTCCTTAGAGCGC 144581  
QY 1118 GACCCGGGATGTGGGGGAGTCTGCGGCTGCGCTGACCAATCGAGTGTGCGTCCATCG 1177  
Db 144582 GCGCGGGATGTGGGGGAGTCTGCGGCTGCGCTGACCAATCGAGTGTGCGTCCATCG 144641  
QY 1178 ACTGCGCTCTGCGAGGCAATTAGGAGCGCTCCCGCGGCGGCTGCGCGCGGCAACCC 1237  
Db 144642 ACTGCGCTCTGCGAGGCAATTAGGAGCGCTCCCGCGGCGGCTGCGCGCGGCAACCC 144701  
QY 1238 AGTGTGTAGTGTGCGGTAGAAACCGTGGCTCTCTGCGCTGAGGCTCTCTGCGCTGAGAG 1297  
Db 144702 GGTGTGTGTGCGGTAGAAACCGTGGCTCTCTGCGCTGAGGCTCTCTGCGCTGAGAG 144761  
QY 1298 GATAAATGCAACGCGGCTATGCACTGGGCTGGGCGCTTTGGGC-----ATCCT 1353  
Db 144762 GGTAAATGCAACGCGGCTATGCACTGGGCTGGGCGCTTTGGGC-----ATCCT 144821  
QY 1354 CCTGCGCTCTAGGGGTTCCAGATCGCCCTTTCTGTTGAGCTGGGAAACACGCGCTG 1413  
Db 144822 CCTGCGCTCTAGGGGTTCCAGATCGCCCTTTCTGTTGAGCTGGGAAACACGCGCTG 144881  
QY 1414 ACTCCAGGACTTGTGTTCTCTCTCACTGGGAAAGGTGGC-GGGGCGAGCTTTTCAAG 1472  
Db 144882 ACTCCAGGACTTGTGTTCTCTCTCACTGGGAAAGGTGGC-GGGGCGAGCTTTTCAAG 144941  
QY 1473 GAGGCGCTGGGAACTTCGAGAGCAGGCTCAGCTCTCACTCTGTGCTCTTAGTTATC 1532

Db	144942	GAAGGGCTCGGGAATTCCGACAGCAGGTACCCCTCTCACTCTGGGGCTGTAGTCTTC	145001
QY	1533	TTGCATGCTCTGGTCTTTTCATACGCTGTCCCTGCACAGGAACCTCCATCCCCATCTT	1592
Db	145002	TTGCATGCTCTGACCTTTTCATAGCGTCTCTCTGCACAGGAACCTCCATCCCCATCTT	145061
QY	1593	TCTCTGCTGTTCGAACCTTCAGAAATCTCAAGGCTCAGCTTAGAGTCACTTCTTCCGA	1652
Db	145062	TTTCTGCTGTTCGAACCTTCAGAAATCTCAAGGCTCAGCTTAGAGTCACTTCTTCCGA	145121
QY	1653	AGCTTTCTCTCAACACCTCCCGCCCTGCTGCTGTGCTGCCCTCAGGCGCTCTCTCACAGC	1712
Db	145122	AGCTTTCTCTCAACACCTCCCGCCCTGCTGCTGTGCTGCCCTCAGGCGCTCTCTCACAC	145181
QY	1713	ACTGATAACAGCTGTCCGTCTCCACCTCCACCACTTCCAATCCCAAGGAGTGA	1772
Db	145182	ACTGATAACAGCTGTCTCTCCACCTCCACCACTTCCAATCCCAAGGAGTGA	145241
QY	1773	GCCACAGAGGGCAGGACAGAGCTGCTGTGTTCTCTGTGTG---CCAGGGCCCCAGCAAG	1829
Db	145242	GCCACAGAGGGCAGGACAGAGCTGCTGTGTTCTCTGTGTG---CCAGGGCCCCAGCAAG	145301
QY	1830	GGAATGTATGGAGGTGGAGGTGCAGGCACTGGGATTAGGGCTTGGGCTCTGGTGT	1889
Db	145302	GGAATGTATGGAGGTGGAGGTGCAGGCACTGGGATTAGGGCTTGGGCTCTGGTGT	145361
QY	1890	TGGAGGCTGGATCTCGATCCCTGTTAGTGGAGTGTCCCTTTACAGCAACTGCGCTGG	1949
Db	145362	CGGA-----AGCTGGATCCCTGTTAGTGGAGTGTCCCTTTACAGTGTGCGCTGG	145415
QY	1950	CTGGCTCGGGCCCTGCTTTTGCTCCTCTTCAGCTGCGGCTGCAGCTGCCA	2000
Db	145416	CTGGCTCGGGCCCTGCTTTTGCTCCTCTTCAGCTGCGGCTGCAGCTGCCA	145466
RESULT 5	AC002472/c		
LOCUS	AC002472	147086 bp	DNA linear PRI 09-APR-2003
DEFINITION	Homo sapiens Chromosome 22q11.2 PAC Clone p_n5 In BCR12-GGT Region, complete sequence.		
ACCESSION	AC002472		
VERSION	AC002472.8	GI:24137490	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Budarf,M.L. and Emanuel,B.S.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 147086)		
AUTHORS	Zhang,G., Zhan,M., Lao,V. and Roe,B.A.		
TITLE	Homo sapiens Chromosome 22q11.2 PAC Clone p_n5 In BCR12-GGT Region		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 147086)		
AUTHORS	Zhang,G., Lao,V., Zhan,M. and Roe,B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-AUG-1997) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
REFERENCE	4 (bases 1 to 147086)		
AUTHORS	Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-SEP-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
REFERENCE	5 (bases 1 to 147086)		
AUTHORS	Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-OCT-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
REFERENCE	6 (bases 1 to 147086)		

FEATURES		Location/Qualifiers
source		1..147086
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="22"
		/map="22q11.2"
		/clone="p_n5"
ORIGIN		
Query Match	51.3%;	Score 1025.4; DB 9; Length 147086;
Best Local Similarity	78.1%;	Pred. No. 3.3e-262;
Matches 1446;	Conservative 0;	Mismatches 236; Indels 169; Gaps 11;
Qy	242	CCTGGCACAATAGCGAGATAAAATAATTATTTAANTTAGCCAGATGTCGTAGCCCCCTGT 301
Db	90979	CGTGACCCACCACCCAGCCAAAATAATTTTATATTAGCAGAGCTGGTGGTCTCTCTGT 90920
Qy	302	AGTCTCAGGACTCAGGAGGCTCAGGCAGGAGGCTCACAGAGTGCAGATTCAAGATG 361
Db	90919	AGTCCCAGCTACTCAGGAGGCTCAGGCAGAGGCTCACTAGAGTGCAGATTCAAGGCTG 90860
Qy	362	CAGTAGCTATGATNCTGCCACTGCATGAAGCTGGGTGACAGAGCAAGACCTTGGCTC 421
Db	90859	CACTGAGCTATTATCTGCCACTGCACCTCAGCCCGGGTGACAGAGCAAGAACCCTGGCTC 90800
Qy	422	TATAAATGAATACATAAAGTCT----- 444
Db	90799	TAATAAATAAATAATTTATTAATTTATTAATAATATGAGCCCAAGAACCC 90740
Qy	445	-----CAGCTAGTG 455
Db	90739	AATGTTTGTGTAGTAGAAGTCATGGCTGGCTGGCTGAGCATCCACAGCTAGTGGTG 90680
Qy	456	GTAGCTAATCTCCAGAGTCAGGCCTCTACCTGCTGATGACAAATGCCACATATGTC 515
Db	90679	GTAGCTAATCTCCAGAGTCAGGCCTCTGGCTGTCTGATGCCAAATTGCATGCTGTGTC 90620
Qy	516	TTTTAACT--GATTGCAGACCACAAATGTTTGTGAATATTTTCCCAGGGAAAAACCG 574
Db	90619	TTTTAACTTGGATTGCAGGCCAATAATGTTTGTAAACATTT--ACGGAGAAAAAAC 90562
Qy	575	GAGTAGTCTCTAAATCTATACATCCATTAATATAGTTTACCTGTGATGGGNAACCC 634
Db	90561	GAGATAGTCTCTAAATCTCATACATCCATTAGCTTAGTTTTCCTGTGATTGGGAAAAAC 90502
Qy	635	CAGCTCTGATTGCATTTCAGGGCGGACAGCCTTTGGTGCATGCTCGCGGATTTTCC 694
Db	90501	CGGCTCTGATTTCAATTCAGGGTGGACAGCCCTTGGTGCATGCTCTGGCGGATTTGC 90442
Qy	695	ATTTTAACCTCTCTTAGAAGCGCCTTCTCATAGTTAAAGTTCTCTGATCGCGCCAGGAGCG 754
Db	90441	ATTTTA-----TCTGATGTTTCCAGGAGCA 90416
Qy	755	CCGAGGAGGGCAGGGGCTGGAGACGCCCGCAGAGGGCTACGTGCCCTGCTCGACAG 814
Db	90415	ATGAGGGAAGGAGGGGGCTGGAGGCGCCCGCAGCGGGTACGTTCCCTGCTGGCCCTT 90356
Qy	815	AGGTCTCTGCCTCCTCGCGGGGCCAGCCACCTCCCAACACCCCTGCGGAGAG--C 872
Db	90355	A-----GCCCCAGCGCGCTCCCCCAACCCCTGCGGAGAGACCC 90316
Qy	873	CCCCAAGGGAGGAGCGGCGCTTGGCCCTTGCCTCCCGAGCACCTTCCGCTCTCTAGTTCGGA 932
Db	90315	CCCAAAGGGAGGAGTCGGGCGCTGCCCTTCCCGCCAGCACCTTCCGCTCTCAGTTCGGA 90256
Qy	933	GTCTGAATCGGCTTGGGACCTCTGCTTGGCTTCGGGGACCCCTGCAAGACGTCACAGGC 992
Db	90255	GTCCCAATTGGCTCGGGAACCTTGCTCGGCTCGGGGACCCCTGAAAGACGCCACAGGC 90196
Qy	993	CGCGCTCGCCTTCTCTCTGCTTTTATNCTCCCGACACCTCTGGCAGGAACCCGCTCAT 1052
Db	90195	CACCATAGCCTTCTTTCTGCTCTTCATCCTCCCGAGACCTCTCACAGGAACCCCTCAT 90136









```

QY 566 AAAAAACCGGAGTAGTCTTAATTTCTATACATCCATTATAGTTTTACCTGTGGATT 625
Db 383 AAAAAACCGGAGTAGTCTTAATTTCTATACATCCATTATAGTTTTACCTGTGGATT 324
QY 626 GGGAAAAACCCAGCTCTGATTGCAATTTTCAGGGCGGGAGCAGCCCTTTGGTGCACCTGTCTGGCG 685
Db 323 GGGAAAAACCCAGCTCTGATTGCAATTTTCAGGGCGGGAGCAGCCCTTTGGTGCACCTGTCTGGCG 264
QY 686 GGATTTTCCATTTTAACCTCTCTAGAACGCGCTTCTCATGTAAAGTTTCTGTATGCCG 745
Db 263 GGATTTTCCATTTTAACCTCTCTAGAACGCGCTTCTCATGTAAAGTTTCTGTATGCCG 204
QY 746 CCAGGAGCGCCGAGGAGAGGGCGGTGGAGAGCGCCCGCAGAGGCTACGTGCCCT 805
Db 203 CCAGGAGCGCCGAGGAGAGGGCGGTGGAGAGCGCCCGCAGAGGCTACGTGCCCT 144
QY 806 GTTGGACAGAGTCTCTCTGCTCTCTGGCGGGCGCAGCCCTCCCAACCCCTGTGGCG 865
Db 143 GTTGGACAGAGTCTCTCTGCTCTCTGGCGGGCGCAGCCCTCCCAACCCCTGTGGCG 84
QY 866 GAGAAGCCCCCAAGGGGAGGAGAGCGGCTGCGCCCTGCGCCGAGCACCTTCGGTCTCTA 925
Db 83 GAGAAGCCCCCAAGGGGAGGAGAGCGGCTGCGCCCTGCGCCGAGCACCTTCGGTCTCTA 24
QY 926 GGTGGAGTCTGAATCGGCTTG 948
Db 23 GGTGGAGTCTGAATCGGCTTG 1

RESULT 9
LOCUS Q142655/c
DEFINITION Sequence 12677 from Patent WO0157276.
ACCESSION Q142655
VERSION Q142655.1 GI:41100027
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human bone marrow
JOURNAL Patent: WO 0157276-A 12677 09-AUG-2001;
FEATURES location/Qualifiers
source 1..563
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/Note="MAP TO AC002472.3-EXPRESSED IN BONE MARROW, SIGNAL = 18"

ORIGIN
Query Match 28.1%; Score 563; DB 6; Length 563;
Best Local Similarity 100.0%; Pred. No. 4.8e-139;
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 CACTGAAAGCTGGGTGACAGAGCAAGACCCCTGGCTCTAATAAATGAATACATAAGTCTC 445
Db 563 CACTGAAAGCTGGGTGACAGAGCAAGACCCCTGGCTCTAATAAATGAATACATAAGTCTC 504
QY 446 ACAGCTAGTGTAGCTAATCTGCGAGAGTCAGGCTCTACCTGTCTGTATGACAAATGGC 505
Db 503 ACAGCTAGTGTAGCTAATCTGCGAGAGTCAGGCTCTACCTGTCTGTATGACAAATGGC 444
QY 506 ACATATGTCTTTTAACTGATTCAGACCAAAATGTTTGTGAATATTTTCCCCAGGG 565
Db 443 ACATATGTCTTTTAACTGATTCAGACCAAAATGTTTGTGAATATTTTCCCCAGGG 384
QY 566 AAAAAACCGGAGTAGTCTTAATTTCTATACATCCATTATAGTTTTACCTGTGGATT 625
Db 383 AAAAAACCGGAGTAGTCTTAATTTCTATACATCCATTATAGTTTTACCTGTGGATT 324

```

```

Db 383 AAAAAACCGGAGTAGTCTTAATTTCTATACATCCATTATAGTTTTACCTGTGGATT 324
QY 626 GGGAAAAACCCAGCTCTGATTGCAATTTTCAGGGCGGGAGCAGCCCTTTGGTGCACCTGTCTGGCG 685
Db 323 GGGAAAAACCCAGCTCTGATTGCAATTTTCAGGGCGGGAGCAGCCCTTTGGTGCACCTGTCTGGCG 264
QY 686 GGATTTTCCATTTTAACCTCTCTAGAACGCGCTTCTCATGTAAAGTTTCTGTATGCCG 745
Db 263 GGATTTTCCATTTTAACCTCTCTAGAACGCGCTTCTCATGTAAAGTTTCTGTATGCCG 204
QY 746 CCAGGAGCGCCGAGGAGAGGGCGGTGGAGAGCGCCCGCAGAGGCTACGTGCCCT 805
Db 203 CCAGGAGCGCCGAGGAGAGGGCGGTGGAGAGCGCCCGCAGAGGCTACGTGCCCT 144
QY 806 GTTGGACAGAGTCTCTCTGCTCTCTGGCGGGCGCAGCCCTCCCAACCCCTGTGGCG 865
Db 143 GTTGGACAGAGTCTCTCTGCTCTCTGGCGGGCGCAGCCCTCCCAACCCCTGTGGCG 84
QY 866 GAGAAGCCCCCAAGGGGAGGAGAGCGGCTGCGCCCTGCGCCGAGCACCTTCGGTCTCTA 925
Db 83 GAGAAGCCCCCAAGGGGAGGAGAGCGGCTGCGCCCTGCGCCGAGCACCTTCGGTCTCTA 24
QY 926 GGTGGAGTCTGAATCGGCTTG 948
Db 23 GGTGGAGTCTGAATCGGCTTG 1

RESULT 10
LOCUS Q178146/c
DEFINITION Sequence 9542 from Patent WO0157274.
ACCESSION Q178146
VERSION Q178146.1 GI:41172885
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human heart
JOURNAL Patent: WO 0157274-A 9542 09-AUG-2001;
FEATURES location/Qualifiers
source 1..563
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/Note="MAP TO AC002472.3-EXPRESSED IN HEART, SIGNAL = 6.7"

ORIGIN
Query Match 28.1%; Score 563; DB 6; Length 563;
Best Local Similarity 100.0%; Pred. No. 4.8e-139;
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 CACTGAAAGCTGGGTGACAGAGCAAGACCCCTGGCTCTAATAAATGAATACATAAGTCTC 445
Db 563 CACTGAAAGCTGGGTGACAGAGCAAGACCCCTGGCTCTAATAAATGAATACATAAGTCTC 504
QY 446 ACAGCTAGTGTAGCTAATCTGCGAGAGTCAGGCTCTACCTGTCTGTATGACAAATGGC 505
Db 503 ACAGCTAGTGTAGCTAATCTGCGAGAGTCAGGCTCTACCTGTCTGTATGACAAATGGC 444
QY 506 ACATATGTCTTTTAACTGATTCAGACCAAAATGTTTGTGAATATTTTCCCCAGGG 565
Db 443 ACATATGTCTTTTAACTGATTCAGACCAAAATGTTTGTGAATATTTTCCCCAGGG 384
QY 566 AAAAAACCGGAGTAGTCTTAATTTCTATACATCCATTATAGTTTTACCTGTGGATT 625
Db 383 AAAAAACCGGAGTAGTCTTAATTTCTATACATCCATTATAGTTTTACCTGTGGATT 324

```

Qy	626	GGGAAACCCAGCTCTGATTGCATTTTCAGGGCGGGACAGCCTTTGGTGCACCTCTCTGGCG	695			
Db	323	GGGAAACCCAGCTCTGATTGCATTTTCAGGGCGGGACAGCCTTTGGTGCACCTCTCTGGCG	264			
Qy	686	GGATTTTCCATTTTAACTCTCTTAGAAGCGCTCTCTCATGTTAAAGTTCCCTGATGCCG	745			
Db	263	GGATTTTCCATTTTAACTCTCTTAGAAGCGCTCTCTCATGTTAAAGTTCCCTGATGCCG	204			
Qy	746	CCAGGAGCCCGAGGAGAGGCGCAGGGGGCTGGAGAGCGCCCGCAGAGGGCTACGTGCCCT	805			
Db	203	CCAGGAGCCCGAGGAGAGGCGCAGGGGGCTGGAGAGCGCCCGCAGAGGGCTACGTGCCCT	144			
Qy	806	GCTGGACAGAGGTCCTCTGCCTCCTCGGCGGGCCGACCCACCTCCCCACAAACCCCTGCGG	865			
Db	143	GCTGGACAGAGGTCCTCTGCCTCCTCGGCGGGCCGACCCACCTCCCCACAAACCCCTGCGG	84			
Qy	866	GAGAAGCCCCCAAGGGGAGAGACGGGCCCTTGGGCCCTTGGCCCGCAGCACCTTCCGTCTCTA	925			
Db	83	GAGAAGCCCCCAAGGGGAGAGACGGGCCCTTGGGCCCTTGGCCCGCAGCACCTTCCGTCTCTA	24			
Qy	926	GFTCGGAGTCTGAATCGGCCCTTG	948			
Db	23	GFTCGGAGTCTGAATCGGCCCTTG	1			
RESULT 11						
CQ225890/c						
LOCUS	CQ225890	563 bp	DNA linear PAT 21-JAN-2004			
DEFINITION	Sequence 12729 from Patent WO0157273.					
ACCESSION	CQ225890					
VERSION	CQ225890.1 GI:41208950					
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.					
AUTHORS	HUMAN GENOME-DEIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO 3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US 60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03 August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000 (03.10.00)<150> US 60/236,359<151> 27 September 2000 (27.09.00)<150> US 60/234,687<151> 21 September 2000 (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>					
JOURNAL	Molecular Dynamics Sequence Listing Patent: WO 0157273-A 12729 09-AUG-2001; Aescima, Inc. (US)					
FEATURES	Location/Qualifiers					
source	1..563 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="MAP TO AC002472.3-EXPRESSED IN ADULT LIVER, SIGNAL = 10"					
ORIGIN						
Query Match 28.1%; Score 563; DB 6; Length 563;						
Best Local Similarity 100.0%; Pred. No. 4.8e-139;						
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	386	CAC TGAAGAGCTGGGTGACAGAGCAGACCTCGGTCTTAATAATGCAATACATAAGTCTC	445			
Db	563	CAC TGAAGAGCTGGGTGACAGAGCAGACCTCGGTCTTAATAATGCAATACATAAGTCTC	504			
Qy	446	ACAGCTAGTGTAGCTAATCTCGCCAGAGTCAGGCTCTACTGCTCATGACAAATGGC	505			
Db	503	ACAGCTAGTGTAGCTAATCTCGCCAGAGTCAGGCTCTACTGCTCATGACAAATGGC	444			
Qy	506	ACACTATGTCCTTTAACTGATTGACAGACCAAAATGTTTGTGTAATATTTTCCCGAGG	565			
Db	443	ACA CTATGTCCTTTAACTGATTGACAGACCAAAATGTTTGTGTAATATTTTCCCGAGG	384			

QY	566	AAAAAACCGGAGTAGTCTTAAATCTCTATACATCCATTATATTAGTTTTACCTGTGGATT	625
Db	383	AAAAAACCGGAGTAGTCTTAAATCTCTATACATCCATTATATTAGTTTTACCTGTGGATT	324
QY	626	GGAAAAACCGAGCTCTGATTTGCATTTTCAGGGGGGACAGCCCTTTGTGTCACTCTCTGGCG	685
Db	323	GGAAAAACCGAGCTCTGATTTGCATTTTCAGGGGGGACAGCCCTTTGTGTCACTCTCTGGCG	264
QY	686	GGATTTTCCATTTTAACCTCTCTCTAGAGCGCCTTCTCATGTGTAAGTTCCCTGATGCCG	745
Db	263	GGATTTTCCATTTTAACCTCTCTCTAGAGCGCCTTCTCATGTGTAAGTTCCCTGATGCCG	204
QY	746	CCAGGAGCGCGAGGAGAGGGGACAGGGGCTTGGAGAGCGCCCGCAGAGGGGCTACGTGCGCT	805
Db	203	CCAGGAGCGCGAGGAGAGGGGACAGGGGCTTGGAGAGCGCCCGCAGAGGGGCTACGTGCGCT	144
QY	806	GCTGGACAGAGGTCCTCTGCTCTCTGGCGGCGCCAGCCACCTCCCAAAACCTCTGGCG	865
Db	143	GCTGGACAGAGGTCCTCTGCTCTCTGGCGGCGCCAGCCACCTCCCAAAACCTCTGGCG	84
QY	866	GAGAAGCCCCCAAGGGAGGAGAGCGGCTTGGCCCTTGGCCCGAGACACCTTCCTGCTCTA	925
Db	83	GAGAAGCCCCCAAGGGAGGAGAGCGGCTTGGCCCTTGGCCCGAGACACCTTCCTGCTCTA	24
QY	926	GGTCGGAGTCTGAATCGGCGCTTG 948	
Db	23	GGTCGGAGTCTGAATCGGCGCTTG 1	
RESULT 12			
CQ263940/c			
LOCUS	CQ263940	563 bp	DNA
DEFINITION	Sequence 12201 from Patent WO0157277.		linear
ACCESSION	CQ263940		
VERSION	CQ263940.1	GI:41236472	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL			
FEATURES			
source			
	1..563		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	/note="MAP TO AC002472.3-EXPRESSED IN FETAL LIVER, SIGNAL		
	= 14"		
ORIGIN			
Query Match	28.1%;	Score 563;	DB 6; Length 563;
Best Local Similarity	100.0%;	Pred. No. 4.8e-139;	
Matches 563;	Conservative	Mismatches 0;	Indels 0; Gaps 0;
QY	386	CACCTCAAGCTGGGTGACAGAGCAAGACCCCTGGCTCTAATAAATGAATACATAAAGTCTC	445
Db	563	CACCTGAAAGCTGGGTGACAGAGCAAGACCCCTGGCTCTAATAAATGAATACATAAAGTCTC	504
QY	446	ACAGCTAGTGTAGCTAATCTGTCCAGAGTCAGGCCCTACCTGTGTGATGACAAATGGC	505
Db	503	ACAGCTAGTGTAGCTAATCTGTCCAGAGTCAGGCCCTACCTGTGTGATGACAAATGGC	444
QY	506	ACACTATGTCCTTTTAAACCTGATTCGACACCACCAATGTTTTGTGAATATTTTCCCCAGGG	565
Db	443	ACACTATGTCCTTTTAAACCTGATTCGACACCACCAATGTTTTGTGAATATTTTCCCCAGGG	384
QY	566	AAAAAACCGGAGTAGTCTTAAATCTCTATACATCCATTATATTAGTTTTACCTGTGGATT	625



QY 686 GGATTTTCCATTTAACTCCTTCTAGAGCGCCTTCTCATGGTAAAGTTCCTGATGCG 745  
Db 263 GGATTTTCCATTTAACTCCTTCTAGAGCGCCTTCTCATGGTAAAGTTCCTGATGCG 204  
QY 746 CCAGAGCGCCAGAGAGGCGAGGGGCTGAGACGCCCGCCGAGAGGCTACGTGCCCT 805  
Db 203 CCAGAGCGCCAGAGAGGCGAGGGGCTGAGACGCCCGCCGAGAGGCTACGTGCCCT 144  
QY 806 GCTGACAGAGTCTCCTGCTCCTCTCGGCGCGCCAGCCCACTCCCAACACCCCTGCGG 865  
Db 143 GCTGACAGAGTCTCCTGCTCCTCTCGGCGCGCCAGCCCACTCCCAACACCCCTGCGG 84  
QY 866 GAGAGCGCCCAAGGGAGGAGACGGGCTGCGCCCTGCGCGCGCCAGACCTTCCGTCCTA 925  
Db 83 GAGAGCGCCCAAGGGAGGAGACGGGCTGCGCCCTGCGCGCGCCAGACCTTCCGTCCTA 24  
QY 926 GGTGCGAGTCTGAATCGGCTTG 948  
Db 23 GGTGCGAGTCTGAATCGGCTTG 1

RESULT 15  
CQ103739/c  
LOCUS 554 bp DNA linear PAT 21-JAN-2004  
DEFINITION Sequence 12598 from Patent WO0157272.  
ACCESSION CQ103739  
VERSION CQ103739.1 GI:41072790  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human placenta  
JOURNAL Patent: WO 0157272-A 12598 09-AUG-2001;  
Neomica, Inc. (US)  
FEATURES  
source  
1. .554  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AC002472.3-EXPRESSED IN PLACENTA, SIGNAL =  
1.2"

ORIGIN  
Query Match 27.7%; Score 554; DB 6; Length 554;  
Best Local Similarity 100.0%; Pred. No. 1.2e-136;  
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 615 ACCTGTGGATTGGAAACCCAGCTCTGATTGCATTTTCAGGGCGGACAGCCCTTTGGTGC 674  
Db 554 ACCTGTGGATTGGAAACCCAGCTCTGATTGCATTTTCAGGGCGGACAGCCCTTTGGTGC 495  
QY 675 ACTGTCTGGCGGATTTTCCATTTTAACTCCTTCTAGAGCGCCTTCTCATGGTAAAGT 734  
Db 494 ACTGTCTGGCGGATTTTCCATTTTAACTCCTTCTAGAGCGCCTTCTCATGGTAAAGT 435  
QY 735 TCCTGATGCCCGCCAGGAGCCCGAGAGAGGGGCTGAGAGCGCCCGCCGAGAGGG 794  
Db 434 TCCTGATGCCCGCCAGGAGCCCGAGAGAGGGGCTGAGAGCGCCCGCCGAGAGGG 375  
QY 795 CTACGTGCCCTGTGGACAGAGGTCTCTGCTCCTCGGCGGGCCAGCCCACTCCAC 854  
Db 374 CTACGTGCCCTGTGGACAGAGGTCTCTGCTCCTCGGCGGGCCAGCCCACTCCAC 315  
QY 855 AACCCCTGCGGGAGAACCCCAAGGGAGGAGCGGCTTGGCCCTGCCCCGAGAC 914  
Db 314 AACCCCTGCGGGAGAACCCCAAGGGAGGAGCGGCTTGGCCCTGCCCCGAGAC 255  
QY 915 TTCCGTCTCTAGTTCGGAGTCTGAATCGGCTTGGGACCCCTGCTTGGCTTCGGGACCCC 974

Db 254 TTCCGTCTCTAGTTCGGAGTCTGAATCGGCTTGGGACCCCTGCTTGGCTTCGGGACCCC 195  
QY 975 TGCAAGAGCTCCACAGGCGCCCTGCTTCCCTCTGCTTCTTATCTTCCCGAGACCT 1034  
Db 194 TGCAAGAGCTCCACAGGCGCCCTGCTTCCCTCTGCTTCTTATCTTCCCGAGACCT 135  
QY 1035 CTGCGAGGAACCGCTCATCGTTACGCCCTTTCGACGCTTCAGACCTTCAGGCGGAGACC 1094  
Db 134 CTGCGAGGAACCGCTCATCGTTACGCCCTTTCGACGCTTCAGACCTTCAGGCGGAGACC 75  
QY 1095 GCTTGGCGCTCACTTACAGCGGACCCGGGAGTGTGGCGGAGTCTGCGGCTGCGCTGA 1154  
Db 74 GCTTGGCGCTCACTTACAGCGGACCCGGGAGTGTGGCGGAGTCTGCGGCTGCGCTGA 15  
QY 1155 CCAATCGAGTGTGG 1168  
Db 14 CCAATCGAGTGTGG 1

Search completed: November 22, 2004, 01:15:03  
Job time : 8566.13 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 6407.8 Seconds  
(without alignments)

11373.559 Million cell updates/sec

Title: US-09-820-095B-3\_COPY\_1\_2000

Perfect score: 2000

Sequence: 1 ttccaagtccatgggtgcc.....agctgggtgcagctgccca 2000

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.\*

1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_est4: \*  
5: gb\_est5: \*  
6: gb\_est6: \*  
7: gb\_est7: \*  
8: gb\_est8: \*  
9: gb\_est9: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	512.4	25.6	1678	3	CR622747 full-length
C 2	492.2	24.6	760	4	BI668929 603294907
C 3	491.8	24.6	841	5	EX336850 BX336850
C 4	483.4	24.2	743	7	CN368433 17006001
C 5	480	24.0	663	6	CD637541 560174650
C 6	479.8	24.0	651	6	CD637276 56017301H
C 7	469.6	23.5	662	6	CD637277 56017301J
C 8	457.4	22.9	586	7	CN368432 170004243
C 9	448.8	22.4	498	6	CD637540 56017465H
C 10	447.4	22.4	646	5	EQ635907 h401310.Y
C 11	429.4	21.5	891	5	EUI31301 AGENCOURT
C 12	421.6	21.1	689	7	CN368430 170004247
C 13	402.2	20.1	416	8	AQ084067 HS 2230.B
C 14	399.8	20.0	573	6	CD637539 56017433J
C 15	399.8	20.0	574	6	CD637543 56017481U
C 16	398.8	20.0	574	6	CD637545 56017489J
C 17	398.2	19.9	574	6	CD637535 56017373J
C 18	397.2	19.9	573	6	CD637542 56017481H
C 19	390.4	19.5	1040	5	EQ696761 AGENCOURT
C 20	387.8	19.4	572	6	CD637544 56017489H
C 21	385.2	19.3	573	6	CD637534 56017373H
C 22	376.8	18.8	568	6	CD637538 56017433H
C 23	371.4	18.6	909	2	BE620507 601483763
C 24	370.4	18.5	615	2	BE391170 601286740

C 25	364.4	18.2	1037	2	BF311025
C 26	345.4	17.3	890	4	BI197333
C 27	345.2	14.9	1004	4	BI197526
C 28	275.2	13.8	473	5	EX280737
C 29	195.6	9.8	481	2	AW418681
C 30	178	8.9	355	6	CD620489
C 31	178	8.9	356	6	CD620488
C 32	178	8.9	357	6	CD620491
C 33	177.4	8.9	379	6	CD620495
C 34	177.4	8.8	346	6	CD620487
C 35	177	8.8	420	6	CD620494
C 36	176.6	8.8	803	2	BE620906
C 37	176.4	8.8	355	6	CD620496
C 38	166	8.3	355	6	CD620490
C 39	162.4	8.1	444	6	CD620493
C 40	135.8	6.8	383	6	CA396428
C 41	130.2	6.5	2945	3	BC028231
C 42	104.6	5.2	537	6	CA439871
C 43	104.2	5.2	725	8	AQ315048
C 44	102.8	5.1	330	6	CD620492
C 45	100	5.0	585	8	AQ540333

## ALIGNMENTS

RESULT 1  
CR622747/c  
LOCUS  
DEFINITION full-length cDNA clone CS0DI033YM19 of Placenta Cot 25-normalized of Homo sapiens (human).  
ACCESSION CR622747  
VERSION CR622747.1 GI:50503554  
KEYWORDS HTC; CNSLT\_cDNA.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1678)  
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue  
REFERENCE 2 (bases 1 to 1678)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
FEATURES  
Location/Qualifiers  
1..1678  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DI033YM19"  
/tissue\_type="Placenta Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"

Query Match 25.6%; Score 512.4; DB 3; Length 1678;  
Best Local Similarity 99.8%; Pred. No. 3.2e-109;  
Matches 513; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 614 TACCTGTGATTGGGAAAAACCCAGCTCTGATTGATTTTCAGGGCGGACAGCTTTGTG 673  
DB 514 TCCCTGTGATTGGGAAAAACCCAGCTCTGATTGATTTTCAGGGCGGACAGCTTTGTG 455

674 CACTGTCTGGCGGATTTTCAATTTAACTCTCTTAGAAGCGCTTCTCATGTAAAG 733  
 454 CACTGTCTGGCGGATTTTCAATTTAACTCTCTTAGAAGCGCTTCTCATGTAAAG 395  
 734 TTCTGTATGTCGCGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 793  
 394 TTCTGTATGTCGCGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 335  
 794 GCTACGTGCTCTGCTGAGACAGAGTCTCTCTGCTCTCTGCGCGCGCGACCTCCCA 853  
 334 GCTACGTGCTCTGCTGAGACAGAGTCTCTGCTCTCTGCGCGCGCGACCTCCCA 275  
 854 CAACCCCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 913  
 274 CAACCCCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 215  
 914 CTTCCCTCTCTAGTCTGAGTCTGATCGGCTCTGAGAGAGAGAGAGAGAGAGAG 973  
 214 CTTCCCTCTCTAGTCTGAGTCTGATCGGCTCTGAGAGAGAGAGAGAGAGAGAG 155  
 974 CTGCAAGAGTCTACAGGCGCGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 1033  
 154 CTGCAAGAGTCTACAGGCGCGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 95  
 1034 TCTGGCAGGACCGCTCATGCTAGCGGCTCTGCGGCTCTGCGGCTCTGCGGCT 1093  
 94 TCTGGCAGGACCGCTCATGCTAGCGGCTCTGCGGCTCTGCGGCTCTGCGGCT 35  
 1094 CGTTGGCGCTCTCACTTAGAGCGGACCGCGGGA 1127  
 34 CGTTGGCGCTCTCACTTAGAGCGGACCGCGGGA 1

## RESULT 2

BI668929/c  
 LOCUS  
 DEFINITION 603294907F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:5314345 5',  
 mRNA sequence.  
 ACCESSION BI668929  
 VERSION  
 KEYWORDS  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM11796 row: m column: 02  
 High quality sequence step: 758.  
 Location/Qualifiers  
 1..760  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5314345"  
 /tissue\_type="hypothalamus"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_96"  
 /note="Organ: brain; Vector: pBluescriptR (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI

## FEATURES

## source

615 ACCTGTGATTGGGAAACCCAGCTCTGATTGCAATTTAGGGCGGACACCTTTGGTGC 674  
 517 ATCTGTGATTGGGAAACCCAGCTCTGATTGCAATTTAGGGCGGACACCTTTGGTGC 458  
 675 ACTGTCTGGCGGATTTTCAATTTAACTCTCTTAGAAGCGCTTCTCATGTAAAGT 734  
 457 ACTGTCTGGCGGATTTTCAATTTAACTCTCTTAGAAGCGCTTCTCATGTAAAGT 398  
 735 TCCTGTATGTCGCGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 794  
 397 TCCTGTATGTCGCGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338  
 795 CTAGTGTCTCTGCTGAGAGAGTCTCTGCTCTCTGCGGCGGCGAGCCCTCCAC 854  
 337 CTAGTGTCTCTGCTGAGAGAGTCTCTGCTCTCTGCGGCGGCGAGCCCTCCAC 278  
 855 AACCCCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 914  
 277 AACCCCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 218  
 915 TTCCGTCTCTGAGTCTGAGTCTGAAATCGGCTCTGAGAGAGAGAGAGAGAGAG 973  
 217 TTCCGTCTCTGAGTCTGAGTCTGAAATCGGCTCTGAGAGAGAGAGAGAGAGAG 158  
 974 CTGCAAGAGTCTACAGGCGCGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 1033  
 157 CTGCAAGAGTCTACAGGCGCGGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 98  
 1034 TCTGGCAGGACCGCTCATGCTAGCGGCTCTGCGGCTCTGCGGCTCTGCGGCT 1093  
 97 TCTGGCAGGACCGCTCATGCTAGCGGCTCTGCGGCTCTGCGGCTCTGCGGCT 38  
 1094 CGTTGGCGCTCTCACTTAGAGCGGACCC 1122  
 37 CGTTGGCGCTCTCACTTAGAGCGGACCC 9

## RESULT 3

## BI668929/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## CONTACT

## Genoscope - Centre National de Sequencage

## Bp 191 91006 EVRY cedex - France

Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr)

## 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime

(gtcggg); Oligo-dT primed using primer  
 5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average  
 insert size 2.3 kb and normalized to ROT 5. This is a  
 primary library enriched for full-length clones and  
 constructed using the Cap-trapper method (Carninci, in  
 preparation). Library constructed by M. Brownstein  
 (NIH/NHGRI, National Institutes of Health). Note: this is  
 a NIH\_MGC Library."

## ORIGIN

Query Match 24.6%; Score 492.2; DB 4; Length 760;  
 Best Local Similarity 99.2%; Pred. No. 1.6e-104;  
 Matches 505; Conservative 0; Mismatches 3; Indels 1; Gaps 1;







ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (Bases 1 to 662)
REFERENCE	Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L. Circular rapid amplification of cDNA ends for high-throughput extension cloning of partial genes
AUTHORS	Genomics 84 (1), 205-210 (2004)
JOURNAL	Contact: Fu GK Incyte Genomics, Inc. 3160 Porter Dr., Palo Alto, CA 94304, USA Tel: 6509454102 Email: gfu@incyte.com.
COMMENT	Location/Qualifiers 1. .662 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone_lib="FLP" /note="Vector: pDrive Cloning Vector"
FEATURES	source

Query Match	23.5%	Score 469.6	DB 6	Length 662
Best Local Similarity	99.0%	Pred. No. 3.3e-99		
Matches 483	Conservative 0	Mismatches 4	Indels 1	Gaps 1
Qy	614	TACCTGTGGATTGGGAAAACCCAGCTCTGATTGCAATTCAGGGCGGACAGCCTTTGGTG	673	
Db	166	TGCTGTGGAVTTGGGAAAACCCAGCTCTGATTGCAATTCAGGGCGGACAGCCTTTGGTG	225	
Qy	674	CACGTGCTGGCGGATTTTCCATTTTAACTTCCTTCTAGAGCGCCTTCTCATGTAAAG	733	
Db	226	CACGTGCTGGCGGATTTTCCATTTTAACTTCCTTCTAGAGCGCCTTCTCATGTAAAG	285	
Qy	734	TTCTGTATCCGCCAGAGAGCGCCAGAGAGAGGGGAGGGGCTGGAGACGCCCGCAGAGG	793	
Db	286	TTCTGTATCCGCCAGAGAGCGCCAGAGAGAGGGGAGGGGCTGGAGACGCCCGCAGAGG	345	
Qy	794	GCTACGTGCCCTGCTGGACAGAGGTCTCTGCCCTCTCGGCGGGCCAGGCCCACTCCCA	853	
Db	346	GCTACGTGCCCTGCTGGACAGAGGTCTCTGCCCTCTCGGCGGGCCAGGCCCACTCCCA	405	
Qy	854	CAACCCCTTGGGGAGAGCCCCCAAGGGGAGGAGACGGGCTGCGCCCTGCCCGCAGCAC	913	
Db	406	CAACCCCTTGGGGAGAGCCCCCAAGGGGAGGAGACGGGCTGCGCCCTGCCCGCAGCAC	465	
Qy	914	CTTCGTCTCTAGTTCGGAGTCTGAATCGGCTTTGGACCTGTCTGGCTTCGGGGACCC	973	
Db	466	CTTCGTCTCTAGTTCGGAGTCTGAATCGGCTTTGGACCTGTCTGGCTTCGGGGACCC	525	
Qy	974	CTGCAAGACGTCACAGSCGCCGTGCGCTCTTCTCTCTGCTTTTATCTCTCCCGACACC	1033	
Db	526	CTGCAAGACGTCACAGSCGCCGTGCGCTCTTCTCTCTGCTTTTATCTCTCCCGACACC	585	
Qy	1034	TCTGGCAGGAACCGCTCATCGTTACGCCCTTTCGCAGCCTCAGACCTCAGGCGGAGAC	1093	
Db	586	TCTGGCAGGAACCGCTCATCGTTATGCCCTTTCGCAGCCTCAGA-CTGAGGCGGAGAC	644	
Qy	1094	CGCTTGGC 1101		
Db	645	CGTTGGGC 652		

RESULT 8	CN368432/c	CN368432	596 bp	linear	EST 16-MAY-2004
LOCUS		1700424376491	GRN_ES	Homo sapiens	cDNA 5', mRNA sequence.
DEFINITION		CN368432			
ACCESSION		CN368432			
VERSION		CN368432.1	GI:47368366		
KEYWORDS		EST.			
SOURCE		Homo sapiens (human)			
ORGANISM		Homo sapiens			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.									
AUTHORS	1 (bases 1 to 586) Bradenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J. and Stanton, L. W.									
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation									
JOURNAL	Nat. Biotechnol. 22 (6): 707-716 (2004)									
COMMENT	Contact: Bradenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbradenberger@geron.com Insert length: 586 Std Error: 0.00.									
FEATURES	Location/Qualifiers source 1..586 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /tissue_type="embryonic stem cells, cell lines H1, H7, and, H9" /clone_lib="GRN ES" /notes="Oligo dT primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"									
ORIGIN	Query Match 22.9%; Score 457.4; DB 7; Length 586; Best Local Similarity 98.5%; Pred. No. 2.4e-96; Matches 472; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  614 TACCTGTGGATTGGGAAACCCCA-GCTCTGATTGCAATTCAGGCGGGAGACGCTTTGGT 672 Db 483 TGCTGTGGATTGGGAAACCCAGGCTCTGATTGCATTCAGGCGGGAGACGCTTTGGT 424  673 GCACCTGTCTGGCGGATTTTCCATTTAACTCTCTTAGAAGCGCTTCTCATGTGTAATA 732 Db 423 GCACCTGTCTGGCGGATTTTCCATTTAACTCTCTTTAGAAGCGCTTCTCATGTGTAATA 364  733 GTTCTCTATGCCCGCAGGAGCGCGAGGAGGCGGAGGCGCTGGAGACGCCCGCCGACAG 792 Db 363 GTTCTCTATGCCCGCAGGAGCGCGAGGAGGCGGAGGCGCTGGAGACGCCCGCCGACAG 304  793 GGCTACCTGTGCCCTGCTGGACAGAGGTCTCTGCTCTCTCGGCGCGCCAGCCACCTCCC 852 Db 303 GGCTACCTGTGCCCTGCTGGACAGAGGTCTCTGCTCTCTCGGCGCGCCAGCCACCTCCC 244  853 ACAACCCCTCGGGAGAAAGCCCAAGGAGAGAGAGCGGCGCTGCCCTGCCCGCCGACGA 912 Db 243 ACAACCCCTCGGGAGAAAGCCCAAGGAGAGAGAGCGGCGCTGCCCTGCCCGCCGACGA 184  913 CATTCCGCTCTAGTTCGGAGTCTGAATCGGCCCTTGGAGCCCTCTTGGCTTCGGGAGACC 972 Db 183 CATTCCGCTCTAGTTCGGAGTCTGAATCGGCCCTTGGAGCCCTCTTGGCTTCGGGAGACC 124  973 CCTGCAAGAGCTCCACAGGCGCGCTCGCCTCTTCTCTGCTTTTATCTCTCCCGACAG 1032 Db 123 CCTGCAAGAGCTCCACAGGCGCGCTCGCCTCTTCTCTGCTTTTATCTCTCCCGACAG 64  1033 CTCTGGCAGGAAACCGCTCATGTTACGCCCTTTTCGAGGCTCTAGACCCCTCAGGCGGAG 1091 Db 63 CTCTGGCAGGAAACCGCTCATGTTACGCCCTTTTCGAGGCTCTAGACCCCTCAGGCGGAG 5									
RESULT 9	CD637540 498 bp mRNA linear EST 12-JAN-2004									
LOCUS	CD637540									
DEFINITION	56017465h1 FLP Homo sapiens cDNA, mRNA sequence.									
ACCESSION	CD637540									
VERSION	CD637540.1 GI:40285551									
KEYWORDS	EST.									

RESULT 9	CD637540	498 bp	linear
LOCUS	56017465H1	FLP Homo sapiens	cdna, mRNA sequence.
DEFINITION	CD637540		
ACCESSION	CD637540.1	GI:40285551	
VERSION	CD637540.1		
KEYWORDS	EST.		



```

Db      149  TTCCGTCCTAGTGGAGTCTGAATCGGCTTGGAGCCCTGTTGGCTTCGGGACCCC 90
Qy      975  TGAAGACGTCACAGCGCCGCTGCTCTTCTCTGCTTTTATCTTCCCAAGCT 1034
Db      89   TGAAGACGTCACAGCGCCGCTGCTCTTCTCTGCTTTTATCTTCCCAAGCT 30
Qy      1035 CTGGCAGGAACCGCTCATCGTTACGCC 1063
Db      29   CTGGCAGGAACCGCTCATCGTTACGCC 1

RESULT 11
LOCUS   BU151301/c
DEFINITION  BU151301 891 bp mRNA linear EST 03-SEP-2002
5', mRNA sequence.
ACCESSION  BU151301
VERSION    BU151301.1 GI:22664833
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 891)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-femail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2538 row: j column: 03
High quality sequence stop: 693.
FEATURES             source
     1..891
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:6341090"
     /tissue_type="neuroblastoma, cell line"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_47"
     /note="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site_2:
     EcoRI; cDNA made by oligo-dT priming. Directionally
     cloned into EcoRI/XhoI sites using the following 5',
     adaptor: GGACGAG(G). Size-selected >500bp for average
     insert size 1.8kb. Library constructed by Ling Hong in
     the laboratory of Gerald M. Rubin (University of
     California, Berkeley) using ZAP-cDNA synthesis kit
     (Stratagene) and Superscript II RT (Life technologies).
     Note: this is a NIH_MGC Library."

ORIGIN
Query Match 21.5%; Score 429.4; DB 5; Length 891;
Best Local Similarity 99.5%; Pred. No. 9.2e-90;
Matches 441; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy      614  TACCTGTGGATTGGGAAACCCAGCTCTGATTTCATTTTCAGGCGGAGCAGCTTTGGTG 673
Db      443  TGCTGTGGATTGGGAAACCCAGCTCTGATTTCATTTTCAGGCGGAGCAGCTTTGGTG 384
Qy      674  CACTGTCTGGCGGATTTTCATTTTAACTCTCTTACGAGCGCCTTCTCATGGTAAAG 733
Db      383  CACTGTCTGGCGGATTTTCATTTTAACTCTCTTACGAGCGCCTTCTCATGGTAAAG 324
Qy      734  TTCTGATGCGGCAGAGCGCGGAGGAGGCGGCTGAGACGCCCGCCGACAGG 793

```

```

Db      323  TTCCTGATGCGCCAGGAGCGCCGAGGAGAGGGGAGCGGGCTGGAGACGCCGCCAGAGG 264
Qy      794  GCTAGTCCCTCTGTCGACAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 853
Db      263  GCTAGTCCCTCTGTCGACAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 204
Qy      854  CAACCCCTTGGCGGAGAAAGCCCAAGGGAGGAGACGGGCTTGGCCCTTGGCCCGGAGCAC 913
Db      203  CAACCCCTTGGCGGAGAGAGCCCAAGGGAGGAGACGGGCTTGGCCCTTGGCCCGGAGCAC 144
Qy      914  CTTCCGTCCTTAGTTCGAGTCTGAATCGGCTTGGAGACCTGCTTGGCTTGGGGAGCCC 973
Db      143  CTTCCGTCCTTAGTTCGAGTCTGAATCGGCTTGGAGACCTGCTTGGCTTGGGGAGCCC 84
Qy      974  CTGCAAGACGTCACAGCGCCGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 1032
Db      83   CTGCAAGACGTCACAGCGCCGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 24
Qy      1033  CTTGGCAGGAACCGCTCATCGT 1055
Db      23   CTTGGCAGGAACCGCTCATCGT 1

RESULT 12
LOCUS   CN368430/c
DEFINITION  CN368430 589 bp mRNA linear EST 16-MAY-2004
17000424702964 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION  CN368430
VERSION    CN368430.1 GI:47368364
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 589)
AUTHORS   Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
JOURNAL   Contact: Brandenberger R
COMMENT   Regenerative Medicine
Geront Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 889 Std Error: 0.00.
Location/Qualifiers
     1..689
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /tissue_type="embryonic stem cells, cell lines H1, H7, and
     H9"
     /clone_lib="GRN_ES"
     /note="oligo dT primed, full-length enriched cDNA library
     from undifferentiated hES cell lines H1 (p32), H7 (p29),
     and H9 (p26) maintained in feeder-free conditions"

FEATURES             source
     1..689
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /tissue_type="embryonic stem cells, cell lines H1, H7, and
     H9"
     /clone_lib="GRN_ES"
     /note="oligo dT primed, full-length enriched cDNA library
     from undifferentiated hES cell lines H1 (p32), H7 (p29),
     and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Query Match 21.1%; Score 421.6; DB 7; Length 689;
Best Local Similarity 99.1%; Pred. No. 6.1e-88;
Matches 424; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      614  TACCTGTGGATTGGGAAACCCAGCTCTGATTTCATTTTCAGGCGGAGCAGCTTTGGTG 673
Db      437  TCCTGTGGATTGGGAAACCCAGCTCTGATTTCATTTTCAGGCGGAGCAGCTTTGGTG 378
Qy      674  CACTGTCTGGCGGATTTTCATTTTAACTCTCTTACGAGCGCCTTCTCATGGTAAAG 733
Db      377  CACTGTCTGGCGGATTTTCATTTTAACTCTCTTACGAGCGCCTTCTCATGGTAAAG 318

```



```

QY 734 TTCTGATGCCCGCAGAGCGCCGAGAGAGGCGGAGGCTGGAGACGCCCGCAGAGG 793
|
|
|
Db 286 TTCTGATGCCCGCAGAGCGCCGAGAGAGGCGGAGGCTGGAGACGCCCGCAGAGG 345
|
|
|
QY 794 GCTACGTGCGCTCTGTCGACAGAGGTCTCTGCGCTCTGCGGCGGCGCAGCCACCTCCCA 853
|
|
|
Db 346 GCTACGTGCGCTCTGTCGACAGAGGTCTCTGCGCTCTGCGGCGGCGCAGCCACCTCCCA 405
|
|
|
QY 854 CAACCCCTCGGGAGAGAGCCCAAGGAGAGAGAGCGGCGCTGGCCCTGCGCCGAGAGAC 913
|
|
|
Db 406 CAACCCCTCGGGAGAGAGCCCAAGGAGAGAGAGCGGCGCTGGCCCTGCGCCGAGAGAC 465
|
|
|
QY 914 CTTCCGCTCTAGGTGCGAGTCTGAATCGGCGCTTGGGACCTGCTTGGCTTGGGAGCCC 973
|
|
|
Db 466 CTTCCGCTCTAGGTGCGAGTCTGAATCGGCGCTTGGGACCTGCTTGGCTTGGGAGCCC 525
|
|
|
QY 974 CTCGAAGACGTCACAGGCGCGCTGCGCTCTTCCTCTGCTT 1016
|
|
|
Db 526 CTCGAAGACGTCACAGGCGCGCTGCGCTCTTCCTCTGCTT 568
|
|
|

RESULT 15
CD637543/c
LOCUS 56017481J1 FLP Homo sapiens cDNA, mRNA linear EST 12-JAN-2004
DEFINITION CD637543
ACCESSION CD637543
VERSION CD637543.1 GI:40285554
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 574)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
JOURNAL extension cloning of partial genes
COMMENT Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com
FEATURES
source
1..574
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 20.0%; Score 399.8; DB 6; Length 574;
Best Local Similarity 99.5%; Pred. No. 7.9e-83;
Matches 401; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 614 TACCTGTGGATTGGGAAACCCAGCTCTGATTCGATTTAGGCGCGGAGAGCGCTTTGGTG 673
|
|
|
Db 408 TGCCTGTGGATTGGGAAACCCAGCTCTGATTCGATTTAGGCGCGGAGAGCGCTTTGGTG 349
|
|
|
QY 674 CACTGTCTGGCGGGAATTTCCATTTAACTCTCTTAGAAGCGCTTCTCATGTAAAG 733
|
|
|
Db 348 CACTGTCTGGCGGGAATTTCCATTTAACTCTCTTAGAAGCGCTTCTCATGTAAAG 289
|
|
|
QY 734 TTCTGTATGCGGCAGAGCGCCGAGAGAGGCGGCGCTGGAGACGCCCGCAGAGG 793
|
|
|
Db 288 TTCCTGATGCGGCAGAGCGCCGAGAGAGGCGGCGCTGGAGACGCCCGCAGAGG 229
|
|
|
QY 794 GCTACGTGCGCTCTGTCGACAGAGTCTCTGCGCTCTGCGGCGGCGCAGCCACCTCCCA 853
|
|
|
Db 228 GCTACGTGCGCTCTGTCGACAGAGTCTCTGCGCTCTGCGGCGGCGCAGCCACCTCCCA 169
|
|
|
QY 854 CAACCCCTCGGGAGAGAGCCCAAGGAGAGAGAGCGGCGCTGCGCCCTGCGCCGAGAGAC 913
|
|
|

```

Search completed: November 21, 2004, 17:35:08  
Job time : 6410.8 secs

```

Db 168 CAACCCCTCGGGAGAGAGCCCAAGGAGAGAGCGGCGCTGCGCCCTGCGCCGAGAGAC 109
|
|
|
QY 914 CTTCCGCTCTCTAGGTGCGAGTCTGAATCGGCGCTTGGGACCCCTGCTTGGCTTCGGGGACCC 973
|
|
|
Db 108 CTTCCGCTCTCTAGGTGCGAGTCTGAATCGGCGCTTGGGACCCCTGCTTGGCTTCGGGGACCC 49
|
|
|
QY 974 CTGCAAGACGTCCTCACAGGCGCGCTGCGCTCTTCTCTCTGCTT 1016
|
|
|
Db 48 CTGCAAGACGTCCTCACAGGCGCGCTGCGCTCTTCTCTCTGCTT 6
|
|
|

```





GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 CompuGen Ltd.  
 OM nucleic - nucleic search, using sw model  
 Run on: November 21, 2004, 17:43:16 ; Search time 988.76 Seconds  
 (without alignments)  
 10930.100 Million cell updates/sec  
 Title: US-09-820-095B-3\_COPY\_1\_2000  
 Perfect score: 2000  
 Sequence: 1 tctccagtcctcatgggccc.....agctggcgctgcagctgccca 2000  
 Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0  
 Searched: 362788 seqs, 2701811610 residues  
 Total number of hits satisfying chosen parameters: 7255776  
 Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Published Applications NA:  
 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
 2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
 6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
 12: /cgn2\_6/ptodata/1/pubpna/US09D\_PUBCOMB.seq.\*  
 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2000	100.0	16449	10	US-09-820-095-3
2	1902.2	95.1	1938	13	US-10-027-632-98169
3	1902.2	95.1	1938	13	US-10-027-632-98170
4	1902.2	95.1	1938	15	US-10-027-632-98169
5	1902.2	95.1	1938	15	US-10-027-632-98170
6	597.6	29.9	598	13	US-10-027-632-41182
7	597.6	29.9	598	15	US-10-027-632-41182
8	563	28.1	563	9	US-09-864-761-9542
9	554	27.7	554	9	US-09-864-761-9446
10	458.4	22.9	1904	15	US-10-094-749-1376
11	383.2	19.2	508	15	US-10-029-386-6681
12	139	7.0	139	9	US-09-864-761-25935

13	138.4	6.9	556	15	US-10-029-386-20399	Sequence 20399, A
14	108.2	5.4	51289	17	US-10-322-281-648	Sequence 648, App
15	103.2	5.2	88191	9	US-09-799-799-3	Sequence 3, Appl1
16	102.2	5.1	1906	13	US-10-027-632-97426	Sequence 97426, A
17	102.2	5.1	1906	15	US-10-027-632-97426	Sequence 97426, A
18	102.2	5.1	304905	16	US-10-271-416-1	Sequence 1, Appl1
19	101.4	5.1	75007	17	US-10-741-601-5612	Sequence 5612, App
20	99.8	5.0	474	13	US-10-027-632-56725	Sequence 56725, A
21	99.8	5.0	474	13	US-10-027-632-56725	Sequence 56725, A
22	99.8	5.0	474	13	US-10-027-632-56725	Sequence 56725, A
23	99.8	5.0	474	13	US-10-027-632-56725	Sequence 56725, A
24	99.8	5.0	474	13	US-10-027-632-56725	Sequence 56725, A
25	99.8	5.0	867	13	US-10-027-632-3763	Sequence 3763, App
26	99	5.0	125534	13	US-10-087-192-1678	Sequence 1678, App
27	97.4	4.9	602	10	US-09-764-891-9533	Sequence 9533, App
28	97.4	4.9	53000	13	US-09-953-611-10	Sequence 10, Appl
29	97.4	4.9	93660	13	US-10-087-192-1384	Sequence 1384, App
30	97.4	4.9	93660	13	US-10-087-192-1390	Sequence 1390, App
31	97.2	4.9	32190	9	US-09-764-869-2209	Sequence 2209, App
32	97.2	4.9	32190	16	US-10-027-577-2209	Sequence 2209, App
33	97.2	4.9	32190	16	US-10-027-577-2209	Sequence 2209, App
34	97	4.9	235070	13	US-10-087-192-1990	Sequence 1990, App
35	96.8	4.8	572	13	US-10-027-632-203457	Sequence 203457, App
36	96.8	4.8	572	13	US-10-027-632-203458	Sequence 203458, App
37	96.8	4.8	572	13	US-10-027-632-203458	Sequence 203458, App
38	96.8	4.8	572	15	US-10-027-632-203457	Sequence 203457, App
39	96.8	4.8	572	15	US-10-027-632-203458	Sequence 203458, App
40	96.8	4.8	572	15	US-10-027-632-203459	Sequence 203459, App
41	96.8	4.8	840	13	US-10-027-632-8044	Sequence 8044, App
42	96.8	4.8	840	15	US-10-027-632-8044	Sequence 8044, App
43	96.6	4.8	27007	9	US-09-764-868-1424	Sequence 1424, App
44	96.4	4.8	6319	9	US-09-595-494-58	Sequence 58, Appl
45	96.2	4.8	167343	9	US-09-962-436-281	Sequence 281, App

ALIGNMENTS

RESULT 1  
 US-09-820-095-3  
 ; Sequence 3, Application US/09820095  
 ; Publication No. US20030233668A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WEI, Ming-Hui et al  
 ; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
 ; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES  
 ; TITLE OF INVENTION: .PROTEINS, AND USES THEREOF  
 ; FILE REFERENCE: CL001202  
 ; CURRENT FILING DATE: 2001-03-29  
 ; CURRENT APPLICATION NUMBER: US/09/820,095  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 16449  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-820-095-3

Query Match	100.0%	Score	2000;	DB	10;	Length	16449;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	2000;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	TCTCCAGTCCATGGTGCCCTGGTAGGAGACAGGGGATGAATGTAACCCCTGCATGCG	60				
Db	1	TCTCCAGTCCATGGTGCCCTGGTAGGAGACAGGGGATGAATGTAACCCCTGCATGCG	60				
QY	61	TATAGCACCTGCTCCCTCCCTCCCTGCTGATCACTACCTGGCCCTATTTTGGCTCTAG	120				
Db	61	TATAGCACCTGCTCCCTCCCTCCCTGCTGATCACTACCTGGCCCTATTTTGGCTCTAG	120				
QY	121	AAGCAGCTTCCTTATGCTCTTAGGACCACTGCCCATATGACATGAACATCGA	180				
Db	121	AAGCAGCTTCCTTATGCTCTTAGGACCACTGCCCATATGACATGAACATCGA	180				

181 GGCTAAGGCAACGCAAACTTTTCTTAAAGTCATACAGCTGTCAAAGAAAGCTGGACA 240  
181 GGCTAAGGCAACGCAAACTTTTCTTAAAGTCATACAGCTGTCAAAGAAAGCTGGACA 240  
241 ACCTGGCAACATACGAGATAAATAATTTAAATTTAGCCAGATGTGATCCGCCCTG 300  
241 ACCTGGCAACATACGAGATAAATAATTTAAATTTAGCCAGATGTGATCCGCCCTG 300  
301 TAGTCTCAGGACTCAGAGGCTGAGGAGAGAGCTCACCAGAGTGAGAGTTCAGAGAT 360  
301 TAGTCTCAGGACTCAGAGGCTGAGGAGAGAGCTCACCAGAGTGAGAGTTCAGAGAT 360  
361 GCAGTGAGCTATGATCTCTGCACTGCACTGAAAGCTGGTGACAGAGCAAGACCTGGCT 420  
361 GCAGTGAGCTATGATCTCTGCACTGCACTGAAAGCTGGTGACAGAGCAAGACCTGGCT 420  
421 CTAAATATGATATCAATAAAGCTCAGAGCTAGTGTAGTATCTCTGAGAGTTCAGGC 480  
421 CTAAATATGATATCAATAAAGCTCAGAGCTAGTGTAGTATCTCTGAGAGTTCAGGC 480  
481 CTCTACCTGTCTGATGACAAATGGCACATATGCTTTTAACTGATTTGACAGACCAAA 540  
481 CTCTACCTGTCTGATGACAAATGGCACATATGCTTTTAACTGATTTGACAGACCAAA 540  
541 TGTTTGTGATATTTTCCAGAGGAAACCGAGTAGTGTCTAAATCTATATATCC 600  
541 TGTTTGTGATATTTTCCAGAGGAAACCGAGTAGTGTCTAAATCTATATATCC 600  
601 APTATATTAGTTTACCTGTGGATGGGAAACCCAGCTCTGATTTGCAATTCAGGCGGG 660  
601 APTATATTAGTTTACCTGTGGATGGGAAACCCAGCTCTGATTTGCAATTCAGGCGGG 660  
661 ACAGCTTTGTGACACTGTCTGGGGGATTTTCCATTTTAACTCTCTAGAGCGCT 720  
661 ACAGCTTTGTGACACTGTCTGGGGGATTTTCCATTTTAACTCTCTAGAGCGCT 720  
721 TCTCATGTAAAGTTCTTGATCCGCGAGAGCGCCGAGGAGAGGCGAGGCGCTGAGA 780  
721 TCTCATGTAAAGTTCTTGATCCGCGAGAGCGCCGAGGAGAGGCGAGGCGCTGAGA 780  
781 GCGCCGCGAGAGGCTAGTCCCTCTGTCGAGAGGCTCTCTGCTCTCCGCGCGCC 840  
781 GCGCCGCGAGAGGCTAGTCCCTCTGTCGAGAGGCTCTCTGCTCTCCGCGCGCC 840  
841 AGCCACCTCCCAACCCCTGCGGAGAGAGCCCAAGGAGGAGAGCGGCGCTGGCC 900  
841 AGCCACCTCCCAACCCCTGCGGAGAGAGCCCAAGGAGGAGAGCGGCGCTGGCC 900  
901 CTGCCCCAGCACTTCCTGCTCTAGTGGAGTCTGAATCGGCTTGGACCTCTTG 960  
901 CTGCCCCAGCACTTCCTGCTCTAGTGGAGTCTGAATCGGCTTGGACCTCTTG 960  
961 GCTTCGGGACCCCTGCAAGACGCTCCACAGCGCGCTCTCTCTCTCTCTTTTA 1020  
961 GCTTCGGGACCCCTGCAAGACGCTCCACAGCGCGCTCTCTCTCTCTCTTTTA 1020  
1021 TCCTCCAGAGCTCTGCGAGAGAGCTGATGTTAGCGCCCTTTTCAGAGCTCAGAC 1080  
1021 TCCTCCAGAGCTCTGCGAGAGAGCTGATGTTAGCGCCCTTTTCAGAGCTCAGAC 1080  
1081 CTGAGGCGAGACCGCTTGGCGCTCACTTAGAGCGGACCCGCGGATGTGGCGGAGTC 1140  
1081 CTGAGGCGAGACCGCTTGGCGCTCACTTAGAGCGGACCCGCGGATGTGGCGGAGTC 1140  
1141 TCGCGCTCTGCTGACCAATCAGTGTGGCTGCTCATGATGAGCTGTGCGAGGCAATTA 1200  
1141 TCGCGCTCTGCTGACCAATCAGTGTGGCTGCTCATGATGAGCTGTGCGAGGCAATTA 1200  
1201 GCGAGCGGCTCTCCCGCGGCGGTGCGCCCGGCAACCCAGTGTAGTGTGCGGTAGAAA 1260  
1201 GCGAGCGGCTCTCCCGCGGCGGTGCGCCCGGCAACCCAGTGTAGTGTGCGGTAGAAA 1260

1261 CCGTGGCTCTCTCGCTGAGGCTCTCTCGCTCAGAGGATAAATCTCAGCGCCACGGGC 1320  
1261 CCGTGGCTCTCTCGCTGAGGCTCTCTCGCTCAGAGGATAAATCTCAGCGCCACGGGC 1320  
1321 TATGACATGGGCTGGGCGCTTGTGGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380  
1321 TATGACATGGGCTGGGCGCTTGTGGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380  
1381 CGCCCCCTTTCTGTTGGACTGSGAAACACGCTCTGACTCCAGGACTTGTGTCTCTCT 1440  
1381 CGCCCCCTTTCTGTTGGACTGSGAAACACGCTCTGACTCCAGGACTTGTGTCTCTCT 1440  
1441 CACTGGGAAAGTGGCGGCGAGCTTTTTCAGAGGCGCTGGGAACTTCGAGAGCCAG 1500  
1441 CACTGGGAAAGTGGCGGCGAGCTTTTTCAGAGGCGCTGGGAACTTCGAGAGCCAG 1500  
1501 GTACCCCTCTCACTCTGTGCTCTTATGTTATCTCTGATCTCTGCTCTCTCTCTCTCT 1560  
1501 GTACCCCTCTCACTCTGTGCTCTTATGTTATCTCTGATCTCTGCTCTCTCTCTCTCT 1560  
1561 CTCCTCTCACAGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1620  
1561 CTCCTCTCACAGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1620  
1621 CAAGGCTCAGCTTAGAGTCACTTCTCTCGGAGCTTCTCTCAACACCTCTCTCTCTCT 1680  
1621 CAAGGCTCAGCTTAGAGTCACTTCTCTCGGAGCTTCTCTCAACACCTCTCTCTCTCT 1680  
1681 CTGCTGTGCTCTCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
1681 CTGCTGTGCTCTCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
1741 CCACACCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800  
1741 CCACACCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800  
1801 TGTTCTCTGTCTGCTCAGGCGCCAGCAAGGAAATGTAGGAGGCTGGAGGTGCAAGGCA 1860  
1801 TGTTCTCTGTCTGCTCAGGCGCCAGCAAGGAAATGTAGGAGGCTGGAGGTGCAAGGCA 1860  
1861 GCTGGGATTAAGGCTTCAAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1920  
1861 GCTGGGATTAAGGCTTCAAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1920  
1921 AAGTGTCTCTTAAACAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980  
1921 AAGTGTCTCTTAAACAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980  
1981 AGCTGGGCTGACGCTGCCA 2000  
1981 AGCTGGGCTGACGCTGCCA 2000

## RESULT 2

US-10-027-632-98169  
; Sequence 98169, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363

		; PRIOR FILING DATE: 1999-11-23		; PRIOR APPLICATION NUMBER: US 60/156,358	
		; PRIOR FILING DATE: 1999-09-28		; PRIOR APPLICATION NUMBER: US 60/146,002	
		; PRIOR FILING DATE: 1999-08-09		; NUMBER OF SEQ ID NOS: 325720	
		; SOFTWARE: FastSeq for Windows Version 4.0		; SEQ ID NO 98169	
		; LENGTH: 1938		; TYPE: DNA	
		; ORGANISM: Human		; US-10-027-632-98169	
		Query Match 95.1%; Score 1902.2; DB 13; Length 1938;			
		Best Local Similarity 99.9%; Pred. No. 0;			
		Matches 1901; Conservative 2; Mismatches 0; Indels 0; Gaps 0;			
QY	1	TTCTCAAGTCCATGGGTGCTGCTAGGAGACAGGGGGATGAAATGTAACCCCTGCTATGCG	60		
DB	36	TTCTCAAGTCCATGGGTGCTGCTAGGAGACAGGGGGATGAAATGTAACCCCTGCTATGCG	95		
QY	61	TATAGCCACCTGCTCCCTCCCTGCTGCATCACTACTGCGCTATTTTGTGCTCTAG	120		
DB	96	TATAGCCACCTGCTCCCTCCCTGCTGCATCACTACTGCGCGCTATTTTGTGCTCTAG	155		
QY	121	AAGCACTGCTTCTATGCTCTTATAGGAGACAGGGGGATGACAGATAAGAACATCGA	180		
DB	156	AAGCACTGCTTCTATGCTCTTATAGGAGACAGGGGGATGACAGATAAGAACATCGA	215		
QY	181	GGCTAAGGCAAGCAAAATCTTTTCTTAAAGTCATACAGCTGTCAAAAGAAAGTGGACA	240		
DB	216	GGCTAAGGCAAGCAAAATCTTTTCTTAAAGTCATACAGCTGTCAAAAGAAAGTGGACA	275		
QY	241	ACCTGGGCAACATACGAGATGATAAAATTTATTTAAATTTAGCCAGATGTTAGCCCTG	300		
DB	276	ACCTGGGCAACATACGAGATGATAAAATTTATTTAAATTTAGCCAGATGTTAGCCCTG	335		
QY	301	TAGTCTCAGGCACTCAGGAGGCTGAGGAGGAGGCTCAACAGAGTGCAGAGTTCAAGGAT	360		
DB	336	TAGTCTCAGGCACTCAGGAGGCTGAGGAGGAGGCTCAACAGAGTGCAGAGTTCAAGGAT	395		
QY	361	CGAGTACATGATCTGCTGCCACTGCACCTGAAAGCTGGGTGACAGAGCAAGACCTGGCT	420		
DB	396	CGAGTACATGATCTGCTGCCACTGCACCTGAAAGCTGGGTGACAGAGCAAGACCTGGCT	455		
QY	421	CTAATAATGATATATAAAGTCTCAGAGCTAGTGGTAGCTTAATCTCGCAGAGTACGC	480		
DB	456	CTAATAATGATATATAAAGTCTCAGAGCTAGTGGTAGCTTAATCTCGCAGAGTACGC	515		
QY	481	CTCTACCTGCTGTGATGACAAATGGCACACTATGCTTTTAACCTGATTGACAGACCAAA	540		
DB	516	CTCTACCTGCTGTGATGACAAATGGCACACTATGCTTTTAACCTGATTGACAGACCAAA	575		
QY	541	TGTTTTGTGAATATTTTCCCGAGGAAACCGGAGTGTCTTAAATTTCTATATATCC	600		
DB	576	TGTTTTGTGAATATTTTCCCGAGGAAACCGGAGTGTCTTAAATTTCTATATATCC	635		
QY	601	ATTATATAGTTTACCTGTGGATTGGGAAAACCCAGCTCTGATTGCAATTTAGGGGGGG	660		
DB	636	ATTATATAGTTTACCTGTGGATTGGGAAAACCCAGCTCTGATTGCAATTTAGGGGGGG	695		
QY	661	ACAGCCTTTGGTGCATGCTGCTGGCGGAAATTTTCATTTTAACTCTCTTCCTAGAGGCT	720		
DB	696	ACAGCCTTTGGTGCATGCTGCTGGCGGAAATTTTCATTTTAACTCTCTTCCTAGAGGCT	755		
QY	721	TCTCATGTAAGTTTCCTGATGCTGCCAGGAGCCCGAGGAGGGGACAGGGGCTGGAGA	780		
DB	756	TCTCATGTAAGTTTCCTGATGCTGCCAGGAGCCCGAGGAGGGGACAGGGGCTGGAGA	815		
QY	781	CGCCCCGAGAGGCTACGTGCCCTGCTGGACAGAGTCTCTGCTCTCTCGCGCGCGCC	840		
DB	816	CGCCCCGAGAGGCTACGTGCCCTGCTGGACAGAGTCTCTGCTCTCTCGCGCGCGCC	875		

QY	841	AGCCCACTCCCAACACCCCTCGCGGAGAAAGCCCAAGGGGAGGAGAGCGGCTCGGCC	900		
DB	876	AGCCCACTCCCAACACCCCTCGCGGAGAAAGCCCAAGGGGAGGAGAGCGGCTCGGCC	935		
QY	901	CTGCCCGGAGACCTTTCCTCTCTAGGTGCGAGTCTGAATCGGCTTGGACCTTGTG	960		
DB	936	CTGCCCGGAGACCTTTCCTCTCTAGGTGCGAGTCTGAATCGGCTTGGACCTTGTG	995		
QY	961	GCTTCGGGACCCCTGCAAGACGTCCACAGGCGCGCTCGGCTCTTCTCTCTCTTTT	1020		
DB	996	GCTTCGGGACCCCTGCAAGACGTCCACAGGCGCGCTCGGCTCTTCTCTCTCTTTT	1055		
QY	1021	TCCTCCCAAGACCTCTGGGAGAAACCGTCTCATCTGTTTACGCCCTTTCCAGCCTCAGAC	1080		
DB	1056	TCCTCCCAAGACCTCTGGGAGAAACCGTCTCATCTGTTTACGCCCTTTCCAGCCTCAGAC	1115		
QY	1081	CTGAGCGGAGACCGCTTGGCGCTCACTTAGAGCGGACCCGGGATGTGGCGGAGTC	1140		
DB	1116	CTGAGCGGAGACCGCTTGGCGCTCACTTAGAGCGGACCCGGGATGTGGCGGAGTC	1175		
QY	1141	TGCGGTGCGCTGAACAAATCGAGTGTGGCTCCATCGACTGGCGTCTGCCACGGCAATTA	1200		
DB	1176	TGCGGTGCGCTGAACAAATCGAGTGTGGCGTCCATCGACTGGCGTCTGCCACGGCAATTA	1235		
QY	1201	GCGAGCGCTCCCCCGCGCGTGCSCCGGCAACCCAGTCTCTAGGTTCGCGTAGAAA	1260		
DB	1236	GCGAGCGCTCCCCCGCGCGTGCSCCGGCAACCCAGTCTCTAGGTTCGCGTAGAAA	1295		
QY	1261	CGGTGCTCTCTGCGCTGAGGCTCTCGCTGAGAGATATAACTGCACGCGCACGGGC	1320		
DB	1296	CGGTGCTCTCTGCGCTGAGGCTCTCGCTGAGAGATATAACTGCACGCGCACGGGC	1355		
QY	1321	TATGCACTGGGCTGGGCGCTTGTGGGATCCTCCCTGCCCTTCCTAGGGGTTCCAGCAT	1380		
DB	1356	TATGCACTGGGCTGGGCGCTTGTGGGATCCTCCCTGCCCTTCCTAGGGGTTCCAGCAT	1415		
QY	1381	CGCCCCCTTTCTGTGACTGGGAAACACGCGCTGACTCCAGSACTTGTGTCTCTACTG	1475		
DB	1416	CGCCCCCTTTCTGTGACTGGGAAACACGCGCTGACTCCAGSACTTGTGTCTCTACTG	1475		
QY	1441	CACTGGGAGAGTGGCGGGGACGCTTTTTCAGGAGGCGCTGGGAACTTCGAGAGCCAG	1500		
DB	1476	CACTGGGAGAGTGGCGGGGACGCTTTTTCAGGAGGCGCTGGGAACTTCGAGAGCCAG	1535		
QY	1501	GTCACTCTCACTCTGTGCTTGTAGTTATCTTGTGATGCTCTGCTTTCGATACGCTG	1560		
DB	1536	GTCACTCTCACTCTGTGCTTGTAGTTATCTTGTGATGCTCTGCTTTCGATACGCTG	1595		
QY	1561	CTCCTGCAACAGAACTCCATCCCATCTTTGTCTGCTTGTTCGAACCTTCAGAAATCTG	1620		
DB	1596	CTCCTGCAACAGAACTCCATCCCATCTTTGTCTGCTTGTTCGAACCTTCAGAAATCTG	1655		
QY	1621	CAAGGTCAGCTTAGAGGTCACTTCTTCCGGAAGCTTTTCTCAACACCTTCCCGCGCTG	1680		
DB	1656	CAAGGTCAGCTTAGAGGTCACTTCTTCCGGAAGCTTTTCTCAACACCTTCCCGCGCTG	1715		
QY	1681	CTGCTGTGCTCAGGCGCTCTCTCAGACACTGATATACAGCTGTCTGCTCTCAACCT	1740		
DB	1716	CTGCTGTGCTCAGGCGCTCTCTCAGACACTGATATACAGCTGTCTGCTCTCAACCT	1775		
QY	1741	CCCAACCTCTCACTCCACCCAGGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG	1800		
DB	1776	CCCAACCTCTCACTCCACCCAGGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG	1835		
QY	1801	TGTTCTCTGTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1860		
DB	1836	TGTTCTCTGTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1895		
QY	1861	GCTGGGATTAGGGGCTTGAAGGCTGGGTGTTTGAAGGCTGGATCT	1903		
DB	1896	GCTGGGATTAGGGGCTTGAAGGCTGGGTGTTTGAAGGCTGGATCT	1938		

QY	841	AGCCACCTCCCAACAACCCCTGCGGGAGAAACCCCAAGGGAGGAGACAGGGCTTGCCCC	900
DB	876	AGCCACCTCCCAACAACCCCTGCGGGAGAAACCCCAAGGGAGGAGACAGGGCTTGCCCC	935
QY	901	CTGCCCCGAGACCTTCCGCTCTAGGTGAGATCTGAATCGGCTTGGACCCCTGCTTG	960
DB	936	CTGCCCCGAGACCTTCCGCTCTAGGTGAGATCTGAATCGGCTTGGACCCCTGCTTG	995
QY	961	GCTTCGGGGACCCCTGCAAGACGTCACAGCGCGCTGCTCTCTCTCTCTCTCTCTCTCT	1020
DB	996	GCTTCGGGGACCCCTGCAAGACGTCACAGCGCGCTGCTCTCTCTCTCTCTCTCTCTCT	1055
QY	1021	TCCTCCCGACCTCTGCGAGGAACCGCTCATCGTTACGCCCTTTGGAGCTCTCAGACC	1080
DB	1056	TCCTCCCGACCTCTGCGAGGAACCGCTCATCGTTACGCCCTTTGGAGCTCTCAGACC	1115
QY	1081	CTGAGGGGAGACCGCTTGGCGCTCACTTAGAGCGGACCCGGGATGTTGGGGGAGTC	1140
DB	1116	CTGAGGGGAGACCGCTTGGCGCTCACTTAGAGCGGACCCGGGATGTTGGGGGAGTC	1175
QY	1141	TGCGGTGCGCTCAACCAATCGAGTGGCGTCCATCGATGCGCTCTGCCACGGCAATTA	1200
DB	1176	TGCGGTGCGCTCAACCAATCGAGTGGCGTCCATCGATGCGCTCTGCCACGGCAATTA	1235
QY	1201	GCAGCGGCTCCCGCGGGGCTGCGCGCGGCAACCGAGTGTGTAGTTCGCTGAGAA	1260
DB	1236	GCAGCGGCTCCCGCGGGGCTGCGCGCGGCAACCGAGTGTGTAGTTCGCTGAGAA	1295
QY	1261	CCGTGCTCTCTCTGCGCTGAGGCTCCCTGCGCTGAGAGGATAAACTGCAGCGCACGGGC	1320
DB	1296	CCGTGCTCTCTCTGCGCTGAGGCTCCCTGCGCTGAGAGGATAAACTGCAGCGCACGGGC	1355
QY	1321	TATGCACTGGCTGGGGGCTTGTGGGCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1380
DB	1356	TATGCACTGGCTGGGGGCTTGTGGGCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1415
QY	1381	CGCCCCCTTTCTGTGACTGGGAAACACGCTGACTCCAGGACTTGTGTCTCTCTCTCT	1440
DB	1416	CGCCCCCTTTCTGTGACTGGGAAACACGCTGACTCCAGGACTTGTGTCTCTCTCTCT	1475
QY	1441	CATGCGGGAAGTGGCGGGGAGCTTTTTCAGAGGGGCTTGGGAACTTCGAGAGCCAG	1500
DB	1476	CATGCGGGAAGTGGCGGGGAGCTTTTTCAGAGGGGCTTGGGAACTTCGAGAGCCAG	1535
QY	1501	GTACCCCTCTCACTCTGTGCTCTTGTATCTTGTGATGCTCTGTCTCTCTCTCTCTCT	1560
DB	1536	GTACCCCTCTCACTCTGTGCTCTTGTATCTTGTGATGCTCTGTCTCTCTCTCTCTCT	1595
QY	1561	CTCCCTGCAACCGAGAACCTCCATCCCATCTTGTCTGCTTGTGAACTTCAGAAATCTG	1620
DB	1596	CTCCCTGCAACCGAGAACCTCCATCCCATCTTGTCTGCTTGTGAACTTCAGAAATCTG	1655
QY	1621	CAAGGTGAGCTTAGAGTCACTTCTTCCGGAAGCTTTCTCTCAACCCCTCCCGCCCTG	1680
DB	1656	CAAGGTGAGCTTAGAGTCACTTCTTCCGGAAGCTTTCTCTCAACCCCTCCCGCCCTG	1715
QY	1681	CTGCTGCTGCCCTCAGGCGCTCTCTCAGCAGCTGATAACAGCTGTCTGCTCTCCACCT	1740
DB	1716	CTGCTGCTGCCCTCAGGCGCTCTCTCAGCAGCTGATAACAGCTGTCTGCTCTCCACCT	1775
QY	1741	CCACACCTCTCACTCCCAACCCAGAGAGTGGCCAGAGGCAAGGACAGAGCTGTCTG	1800
DB	1776	CCACACCTCTCACTCCCAACCCAGAGAGTGGCCAGAGGCAAGGACAGAGCTGTCTG	1835
QY	1801	TGTTCTCTGTGTCAGGGGCGGAGCAAGGGGATGTTAGGGGCTGAGGAGTGGAGGCA	1860
DB	1836	TGTTCTCTGTGTCAGGGGCGGAGCAAGGGGATGTTAGGGGCTGAGGAGTGGAGGCA	1895
QY	1861	GCTGGGATAGGGGTTGAGGCTGGGTTGAGGCTGGATCT 1903	
DB	1896	GCTGGGATAGGGGTTGAGGCTGGGTTGAGGCTGGATCT 1938	

```
RESULT 3
US-10-027-632-98170
; Sequence 98170, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98170
; LENGTH: 1938
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-98170

Query Match          95.1%; Score 1902.2; DB 13; Length 1938;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1901; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTCAAGTCATGGTGCCTCCCTGCGATCACTACCTGGGCTATTTTTGGCTCTAG 60
Db 36 TCTCAAGTCATGGTGCCTCCCTGCGATCACTACCTGGGCTATTTTTGGCTCTAG 95
QY 61 TATAGCCACCTGCTCCCTGCGATCACTACCTGGGCTATTTTTGGCTCTAG 120
Db 96 TATAGCCACCTGCTCCCTGCGATCACTACCTGGGCTATTTTTGGCTCTAG 155
QY 121 AAGCACTGCTTCTATGCTCTTAAAGTCATAGCTGCTCAAGATGAAGAACATCGA 180
Db 156 AAGCACTGCTTCTATGCTCTTAAAGTCATAGCTGCTCAAGATGAAGAACATCGA 215
QY 181 GGCTAAGCAACGCAATCTTTTCTTAAAGTCATAGCTGCTCAAGATGAAGAACATCGA 240
Db 216 GGCTAAGCAACGCAATCTTTTCTTAAAGTCATAGCTGCTCAAGATGAAGAACATCGA 275
QY 241 ACCTGGGCAACATAGCGAGATAAATAATTTAAATTAAGCGAGATGTTAGCCCTG 300
Db 276 ACCTGGGCAACATAGCGAGATAAATAATTTAAATTAAGCGAGATGTTAGCCCTG 335
QY 301 TAGTCTAGCACTCAGAGGCTCAGGAGGCTCAGGAGGCTCAGGAGGCTCAGGAGG 360
Db 336 TAGTCTAGCACTCAGAGGCTCAGGAGGCTCAGGAGGCTCAGGAGGCTCAGGAGG 395
QY 361 GCAGTACCTATGATCTCTGCACTGCACTGAAAGCTGGGTGACAGAACACCTGGCT 420
Db 396 GCAGTACCTATGATCTCTGCACTGCACTGAAAGCTGGGTGACAGAACACCTGGCT 455
QY 421 CTAATAAATGAATACATAAAGTCTCAGAGTGTAGTGTAGTGTAGTGTAGTGTAG 480
Db 456 CTAATAAATGAATACATAAAGTCTCAGAGTGTAGTGTAGTGTAGTGTAGTGTAG 515
QY 481 CTCTACCTGTCTGATGACAAATGGCACAATGTCTTTTAACTGTATGACAGACAAA 540
Db 516 CTCTACCTGTCTGATGACAAATGGCACAATGTCTTTTAACTGTATGACAGACAAA 575
```

Db 1656 CAAGGGTACGCTTAGAGGTCACTTCTCGGAGGCTTTCTCAACACACCCCTCCCGCCCTG 1715  
Qy 1681 CTGCTGCTGCTCAGGCTCTCTCTCACAGCACTGATAAAGCTGTCTCGCTCTCCACCT 1740  
Db 1716 CTGCTGCTGCTCAGGCTCTCTCTCACAGCACTGATAAAGCTGTCTCGCTCTCCACCT 1775  
Qy 1741 CCCACCACTCCACTCCCAACCCAGGAAGTGAGGCGCAGAGGCGCAGAGCTGTCTGC 1800  
Db 1776 CCCACCACTCCACTCCCAACCCAGGAAGTGAGGCGCAGAGGCGCAGAGCTGTCTGC 1835  
Qy 1801 TGTCTCTGTGTGCGCAGGCGCCAGCAAGGGAATGTAGGAGGCTGGAGGTGCAAGGCA 1860  
Db 1836 TGTCTCTGTGTGCGCAGGCGCCAGCAAGGGAATGTAGGAGGCTGGAGGTGCAAGGCA 1895  
Qy 1861 GTTGGGATPAGGGGTGAGGGGTGGGTGTTGGAGGCTGGATCT 1903  
Db 1896 GCTGGGATPAGGGGTGAGGGGTGGGTGTTGGAGGCTGGATCT 1938

RESULT 4  
US-10-027-632-98169  
; Sequence 98169, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108927.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 98169  
; LENGTH: 1938  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-98169

Query Match 95.1%; Score 1902.2; DB 15; Length 1938;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1901; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTCAAGTCCATGGGTGCTGCTGGTGGAGACAGGGGATGAATGTGAACCCCTGCAATGGC 60  
Db 36 TCTCAAGTCCATGGGTGCTGCTGGTGGAGACAGGGGATGAATGTGAACCCCTGCAATGGC 95  
Qy 61 TATAGCCACCTGCTCTCTCCCTGCCCTGCATCACTACCTGGCCCTATTTTTGGCTCTAG 120  
Db 96 TATAGCCACCTGCTCTCTCCCTGCCCTGCATCACTACCTGGCCCTATTTTTGGCTCTAG 155  
Qy 121 AAGCACTGCTTCTCTATGCTCTTAGGACCACTGCGCCGATATGACAGATGAACATCGA 180  
Db 156 AAGCACTGCTTCTCTATGCTCTTAGGACCACTGCGCCGATATGACAGATGAACATCGA 215  
Qy 181 GGCTAAGGCAACGCAAAATCTTTCTTAAAGTCATACAGCTGTCAAAAGAAAGCTGGACA 240  
Db 216 GGCTAAGGCAACGCAAAATCTTTCTTAAAGTCATACAGCTGTCAAAAGAAAGCTGGACA 275

Qy 241 ACCTGGGCAACATAGCGAGATAAAAAATATTTAAATTAGCCAGATGCTGTAGCCCTG 300  
Db 276 ACCTGGGCAACATAGCGAGATAAAAAATATTTAAATTAGCCAGATGCTGTAGCCCTG 335  
Qy 301 TAGTCTCAGGACCTCAGGAGGCTGAGCGAGAGGCTCACAGAGTGCAGAGTTCAAGGAT 360  
Db 336 TAGTCTCAGGACCTCAGGAGGCTCAGGAGGAGGCTCACAGAGTGCAGAGTTCAAGGAT 395  
Qy 361 GCAGTGAAGCTATGATCTGCTGCACTGCACTGAAAGCTGGGTGACAGAGCAAGACCTGGCT 420  
Db 396 GCAGTGAAGCTATGATCTGCTGCACTGCACTGAAAGCTGGGTGACAGAGCAAGACCTGGCT 455  
Qy 421 CTAAATAATGAATACATAAAAGTCTCACAGCTAGTGGTAGTAATCTCTGCCAGAGTCAAGC 480  
Db 456 CTAAATAATGAATACATAAAAGTCTCACAGCTAGTGGTAGTAATCTCTGCCAGAGTCAAGC 515  
Qy 481 CTCTACCTGTCTGATGACAAATGGCACATATGCTTTTAACTCTGATGTGAGACCAAA 540  
Db 516 CTCTACCTGTCTGATGACAAATGGCACATATGCTTTTAACTCTGATGTGAGACCAAA 575  
Qy 541 TGTTTTGTGAATATTTTCCCGCAGGAAAAAACCGGAAGTAGTTCTAAATTTCTATACATCC 600  
Db 576 TGTTTTGTGAATATTTTCCCGCAGGAAAAAACCGGAAGTAGTTCTAAATTTCTATACATCC 635  
Qy 601 ATTATATAGTTTAACTGTGGATTTGGGAAAAACCCAGCTCTGATTTGATTTTCAGGCGG 660  
Db 636 ATTATATAGTTTAACTGTGGATTTGGGAAAAACCCAGCTCTGATTTGATTTTCAGGCGG 695  
Qy 661 ACAGCCTTTTGGTCACTGTCTGGGGATTTTCCATTTTAACTCTCTTAGAGGCTCT 720  
Db 696 ACAGCCTTTTGGTCACTGTCTGGGGATTTTCCATTTTAACTCTCTTAGAGGCTCT 755  
Qy 721 TCTCATGTAAAGTTTCTGTATGCGCGCAGGAGCGCGAGAGGAGGAGGAGGAGGAGGAGG 780  
Db 756 TCTCATGTAAAGTTTCTGTATGCGCGCAGGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGG 815  
Qy 781 CGCCCGCAGAGGCTACGTGCGCTCTCTGACAGAGTCTCTGCTCTCTGCGCGCGCC 840  
Db 816 CGCCCGCAGAGGCTACGTGCGCTCTCTGACAGAGTCTCTGCTCTCTGCGCGCGCC 875  
Qy 841 AGCCCACTCCCAACACCCCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900  
Db 876 AGCCCACTCCCAACACCCCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 935  
Qy 901 CTGCGCGGAGGAGGCTCTGCTGCTAGTGGAGTCTGAATCGGCTCTGGAACCTGCTG 960  
Db 936 CTGCGCGGAGGAGGCTCTGCTGCTAGTGGAGTCTGAATCGGCTCTGGAACCTGCTG 995  
Qy 961 GCTTCGGGAGCCCTGCGAGAGGCTCCACAGGCGCGGCTGCTCTCTCTCTCTCTTTTA 1020  
Db 996 GCTTCGGGAGCCCTGCGAGAGGCTCCACAGGCGCGGCTGCTCTCTCTCTCTCTTTTA 1055  
Qy 1021 TCTTCCCGAGCCTCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080  
Db 1056 TCTTCCCGAGCCTCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1115  
Qy 1081 CTGAGCGGAGGAGGCTGCGCGCTCACTTAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
Db 1116 CTGAGCGGAGGAGGCTGCGCGCTCACTTAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1175  
Qy 1141 TCGCGGCTGCGCTGAGCAATTCAGAGTGTGGCTCCATCGATGCGGCTGTGCGAGCAATTA 1200  
Db 1176 TCGCGGCTGCGCTGAGCAATTCAGAGTGTGGCTCCATCGATGCGGCTGTGCGAGCAATTA 1235  
Qy 1201 GCGAGCGGCTCTCCCGCGGCGGCTGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
Db 1236 GCGAGCGGCTCTCCCGCGGCGGCTGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1295  
Qy 1261 CCGTGGCTCTCTGCGGCTGAGGCTCTCTGCGCTGAGAGGATGAACCTGCGCGGAGGAGG 1320  
Db 1296 CCGTGGCTCTCTGCGGCTGAGGCTCTCTGCGCTGAGAGGATGAACCTGCGCGGAGGAGG 1355  
Qy 1321 TATGCACTGGGCTGGGCGCTTTGTGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380

1356 TATGACTGGCTGGGCGCTTGTGGGCATCTCCCTGCTTCTTAGGGGTTCCAGCAT 1415  
1381 CGCCCCCTTTCTGGAAGTGGGAAACACGCTCAGCTCCAGGACTTGTGTGCTCCTCAGTG 1440  
1416 CGCCCCCTTTCTGGAAGTGGGAAACACGCTCAGCTCCAGGACTTGTGTGCTCCTCAGTG 1475  
1441 CACTGGGAAAGTGGGCGGCGGAGCTTTTCAGAGGGGCTTGGGAACTTCGAGAGCCAG 1500  
1476 CACTGGGAAAGTGGGCGGCGGAGCTTTTCAGAGGGGCTTGGGAACTTCGAGAGCCAG 1535  
1501 GTACCCCTCTCACTCTGTGCTCTTAGTTATCTTGATCTCTGCTCTGCTCTGCTCTGCTCTG 1560  
1536 GTACCCCTCTCACTCTGTGCTCTTAGTTATCTTGATCTCTGCTCTGCTCTGCTCTGCTCTG 1595  
1561 CTCCTCTCACAGGAACTCCATCCCACTCTTGTCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 1620  
1596 CTCCTCTCACAGGAACTCCATCCCACTCTTGTCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 1655  
1621 CAAGGGTCAGCTTAGAGGTCACTTCTTCGGAAGCTTCTTCAACACCTTCCCGGCGCTG 1680  
1656 CAAGGGTCAGCTTAGAGGTCACTTCTTCGGAAGCTTCTTCAACACCTTCCCGGCGCTG 1715  
1681 CTGCTGTGCTCAGGCGCTCTCTCTCACAGCACTGATACAGCTGTCTGCTCTTCCACCTT 1740  
1716 CTGCTGTGCTCAGGCGCTCTCTCTCACAGCACTGATACAGCTGTCTGCTCTTCCACCTT 1775  
1741 CCCACCACTCTCACTCCCACTCCAGGAGTGAAGCCAGAGGCGGAGGAGGAGGAGGAGGAGGAG 1800  
1776 CCCACCACTCTCACTCCCACTCCAGGAGTGAAGCCAGAGGCGGAGGAGGAGGAGGAGGAGGAG 1835  
1801 TGTTCTCTGTGTCAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860  
1836 TGTTCTCTGTGTCAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1895  
1861 GCTGGGATAGGGGTTAGGGGCTGGGTGTTGGAGGCTGGATCT 1903  
1896 GCTGGGATAGGGGTTAGGGGCTGGGTGTTGGAGGCTGGATCT 1938

RESULT 5  
US-10-027-632-98170  
; Sequence 98170, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 98170  
; LENGTH: 1938  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-98170

Query Match 95.1%; Score 1902.2; DB 15; Length 1938;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1901; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTCAAGTCCATGGGTGCTGTGTAGGAGACAGGGGATCAATGTGTGAAACCCCTGCTGATGGC 60  
DB 36 TCTCAAGTCCATGGGTGCTGTGTAGGAGACAGGGGATCAATGTGTGAAACCCCTGCTGATGGC 95  
QY 61 TATAGCCACCTGCTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
DB 96 TATAGCCACCTGCTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155  
QY 121 AAGCACTGCTTCTCTATGCTCTCTTAGGACCACTGCGCCGATATGACAGATAGCAATAGCAATCGA 180  
DB 156 AAGCACTGCTTCTCTATGCTCTCTTAGGACCACTGCGCCGATATGACAGATAGCAATAGCAATCGA 215  
QY 181 GGTAAAGCAACGCAAAATCTTTTCTTAAAGTCAATACAGCTGTCAAAAGAAAGGAGTGCACA 240  
DB 216 GGTAAAGCAACGCAAAATCTTTTCTTAAAGTCAATACAGCTGTCAAAAGAAAGGAGTGCACA 275  
QY 241 ACCTGGGCAACATAGCAGATATAAAATTTTAAATTTAGCCAGATGTGTGTAGGCGCCCTG 300  
DB 276 ACCTGGGCAACATAGCAGATATAAAATTTTAAATTTAGCCAGATGTGTGTAGGCGCCCTG 335  
QY 301 TAGTCTCAGGCACTCAGGAGGCTGAGGAGGCTGACAGAGTGCAGAGTGCAGAGTTCAGAGAT 360  
DB 336 TAGTCTCAGGCACTCAGGAGGCTGAGGAGGCTGACAGAGTGCAGAGTTCAGAGAT 395  
QY 361 CGAGTGAAGTATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 396 CGAGTGAAGTATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455  
QY 421 CTAAATAAATGAATACATAAAGTCTCACAGCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480  
DB 456 CTAAATAAATGAATACATAAAGTCTCACAGCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 515  
QY 481 CTCTACCTGTCTGATGACAAATGCGACACTATGCTTTTACCTGATTTGACAGACCAAA 540  
DB 516 CTCTACCTGTCTGATGACAAATGCGACACTATGCTTTTAACTGATTTGACAGACCAAA 575  
QY 541 TGTTTTGTGAATATTTTCCCGAGGAAACCCGAGAGTAGTCTTAAATTTCTATACATCC 600  
DB 576 TGTTTTGTGAATATTTTCCCGAGGAAACCCGAGAGTAGTCTTAAATTTCTATACATCC 635  
QY 601 ATTATATAGTTTACCTGTGATTTGGGAAACCCAGCTCTGATGTGATTTTCAGGCGGG 660  
DB 636 ATTATATAGTTTACCTGTGATTTGGGAAACCCAGCTCTGATGTGATTTTCAGGCGGG 695  
QY 661 ACAGGCTTTGGTGCACCTGCTGGGCGGATTTTCAATTTTAACTCTCTCTAGAGAGCGCT 720  
DB 696 ACAGGCTTTGGTGCACCTGCTGGGCGGATTTTCAATTTTAACTCTCTCTAGAGAGCGCT 755  
QY 721 TCTCATGTGTAAAGTTCTCTGATGCTCCCGAGGAGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
DB 756 TCTCATGTGTAAAGTTCTCTGATGCTCCCGAGGAGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 815  
QY 781 CGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
DB 816 CGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875  
QY 841 AGCCCACTCTCCCAACAAACCCCTGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
DB 876 AGCCCACTCTCCCAACAAACCCCTGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 935  
QY 901 CTGCCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
DB 936 CTGCCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 995  
QY 961 GCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
DB 996 GCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1055  
QY 1021 TCTTCCCCAGACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080



Db	1056	TCCTCCCCAGACCTCTGGCAGGAACCGCTCATCGTTACGCCCTTTTCGAGCCTCAGACC	1115
Qy	1081	CTGAGGCGGAGACCGCTTTGGCGCCTCACTTAGACGCGAACCCGGGGATGTGGCGGAGTC	1140
Db	1116	CTGAGGCGGAGACCGCTTTGGCGCCTCACTTAGACGCGAACCCGGGGATGTGGCGGAGTC	1175
Qy	1141	TGCGGCTGCGCTGACCAATCAGATGTGGCGTCCATCGACTGGCGTCTGCCACGCGCAATTA	1200
Db	1176	TGCGGCTGCGCTGACCAATCGATGTGGCGTCCATCGACTGGCGTCTGCCACGCGCAATTA	1235
Qy	1201	GGGACGGCTCCCGCGGGCGGTTCGCCCGCGGCAACCCAGTGTGTAGTTTGCCTGAGAA	1260
Db	1236	GGGACGGCTCCCGCGGGCGGTTCGCCCGCGGCAACCCAGTGTGTAGTTTGCCTGAGAA	1295
Qy	1261	CCGTGGCTCTCCTCGCTGAGGCTCCCGCTCAGAGGATAAATCGCACGCGCCACGGGC	1320
Db	1296	CCGTGGCTCTCCTCGCTGAGGCTCTCCTCGCTGAGAGATAAATCGCACGCGCCACGGGC	1355
Qy	1321	TATGCACTGGGCTGGCGCCTTTGGGCAATCCCTCCCTGCTTCTAGGGGTTCCAGCAT	1380
Db	1356	TATGCACTGGGCTGGCGCCTTTGGGCAATCCCTCCCTGCTTCTAGGGGTTCCAGCAT	1415
Qy	1381	CGCCCCCTTTCTGTTGACTGGGAAACACGCGCTGACTCCAGGACTGTGTGTCTCTCACTG	1440
Db	1416	CGCCCCCTTTCTGTTGACTGGGAAACACGCGCTGACTCCAGGACTGTGTGTCTCTCACTG	1475
Qy	1441	CATGGGGAAGTGGCGGGGCGAGCTTTTCAGAGGGCCTGGGGAATCTCGCAGAGCCAG	1500
Db	1476	CATGGGGAAGTGGCGGGGCGAGCTTTTCAGAGGGCCTGGGGAATCTCGCAGAGCCAG	1535
Qy	1501	GTCACTCTCACTCTGTGCGCTCTTAGTTATCTTGATGCTCTGTGCTTTTGATACGCTG	1560
Db	1536	GTCACTCTCTCACTCTGTGCGCTCTTAGTTATCTTGATGCTCTGTGCTTTTGATACGCTG	1595
Qy	1561	CTCCCTGCACACGGAACCTCCATCCCATCTTTGTCTGCTTTGTGCGAACTTCAGAAATCTG	1620
Db	1596	CTCCCTGCACACGGAACCTCCATCCCATCTTTGTCTGCTTTGTGCGAACTTCAGAAATCTG	1655
Qy	1621	CAGGGTCAGCTTAGAGGTCACTTTTCGGGAAGCTTTCTCAACACCTTCCCGCCCTG	1680
Db	1656	CAGGGTCAGCTTAGAGGTCACTTTTCGGGAAGCTTTCTCAACACCTTCCCGCCCTG	1715
Qy	1681	CTGCTGTGCGCTCAGGCGCTCTCTCAGAGCACTGATAACAGCTCTGCGTCTCCACCT	1740
Db	1716	CTGCTGTGCGCTCAGGCGCTCTCTCAGAGCACTGATAACAGCTCTGCGTCTCCACCT	1775
Qy	1741	CCACCACTCTCACTCCACCCGAGGAATGAGGCCAGAGGGCAGGGAACAGAGTCTCTGC	1800
Db	1776	CCACCACTCTCACTCCACCCGAGGAATGAGGCCAGAGGGCAGGGAACAGAGTCTCTGC	1835
Qy	1801	TGTTCTCTGTGCCAGGGCCAGCAAAAGGAATGTAGGAGGGTGGGAGGTGCAGGGCA	1860
Db	1836	TGTTCTCTGTGCCAGGGCCAGCAAAAGGAATGTAGGAGGGTGGGAGGTGCAGGGCA	1895
Qy	1861	GCTGGGATTAGGGGTTGAGGGCTGGGTGTGAGAGGCTGATCT	1903
Db	1896	GCTGGGATTAGGGGTTGAGGGCTGGGTGTGAGAGGCTGATCT	1938

## RESULT 6

```

US-10-027-632-41182
; Sequence 41182, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

```

```

, PRIOR APPLICATION NUMBER: US 60/198,676
, PRIOR FILING DATE: 2000-04-20
, PRIOR APPLICATION NUMBER: US 60/193,483
, PRIOR FILING DATE: 2000-03-29
, PRIOR APPLICATION NUMBER: US 60/185,218
, PRIOR FILING DATE: 2000-02-24
, PRIOR APPLICATION NUMBER: US 60/167,363
, PRIOR FILING DATE: 1999-11-23
, PRIOR APPLICATION NUMBER: US 60/156,358
, PRIOR FILING DATE: 1999-09-28
, PRIOR APPLICATION NUMBER: US 60/146,002
, PRIOR FILING DATE: 1999-08-09
, NUMBER OF SEQ ID NOS: 325720
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 41182
, LENGTH: 598
, TYPE: DNA
, ORGANISM: Human
US-10-027-632-41182

```

Query Match 29.9%; Score 597.6; DB 13; Length 598;  
Best Local Similarity 99.8%; Pred. No. 1.2e-173;  
Matches 597; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	116	TCTAGAAGCACTGCTTCTCTATGCTCTTCTAGGACCACTGCCCGCATATGACAGATAAGAAC	175
Db	1	TCTAGAAGCACTGCTTCTCTATGCTCTTCTAGGACCACTGCCCGCATATGACAGATAAGAAC	60
Qy	176	ATCGAGGCTAAGCAAGCGMAATCTTTTCTTAAAGTCATACAGCTGTCTAAAAGAAAGCT	235
Db	61	ATCGAGGCTAAGCAAGCGMAATCTTTTCTTAAAGTCATACAGCTGTCTAAAAGAAAGCT	120
Qy	236	GGCAACACCTGGGCAACATATAGCGAGATAAAAAATTTAAATTAGCCAGATGGGTAGCC	295
Db	121	GGCAACACCTGGGCAACATATAGCGAGATAAAAAATTTAAATTAGCCAGATGGGTAGCC	180
Qy	296	CCCTGTAGTCTCAGCGACTCAGGAGGCTCAGGAGGAGGCTCACAGAGTGCAGAGTTCA	355
Db	181	CCCTGTAGTCTCAGCGACTCAGGAGGCTCAGGAGGAGGCTCACAGAGTGCAGAGTTCA	240
Qy	356	AGGATCGAGTGAGCTATGATCTGCCACCTGCACCTGAAAGCTGGGTGACAGAGCAAGACC	415
Db	241	AGGATCGAGTGAGCTATGATCTGCCACCTGCACCTGAAAGCTGGGTGACAGAGCAAGACC	300
Qy	416	TGGCTCTAATAATGAATACATAAAGTCTCACAGCTAGTGGTAGCTAATCTCCCGAGAGT	475
Db	301	TGGCTCTAATAATGAATACATAAGTCTCACAGCTAGTGGTAGCTAATCTCCCGAGAGT	360
Qy	476	CAGGCCTCTACCTGCTCTGATGACAAATGGCACACTATGCTTTTAAACCTGATTCGAGACC	535
Db	361	CAGGCCTCTACCTGCTCTGATGACAAATGGCACACTATGCTTTTAAACCTGATTCGAGACC	420
Qy	536	ACAAATGTTTTCTGAAATATTTTCCCGAGGAAAAAACCGAAGTAGTTCTTAATTTCTATA	595
Db	421	ACAAATGTTTTCTGAAATATTTTCCCGAGGAAAAAACCGAAGTAGTTCTTAATTTCTATA	480
Qy	596	CATCCATTATATAGTTTTACCTGTGGATTGGGAAAAACCCAGCTCTGATTCGATTTTCAGG	655
Db	481	CATCCATTATATAGTTTTACCTGTGGATTGGGAAAAACCCAGCTCTGATTCGATTTTCAGG	540
Qy	656	GGGGACAGCCTTTGGTGACCTGTCTGGCGGATTTTCCATTTTAAACCTCTCTTAGA	713
Db	541	GGGGACAGCCTTTGGTGACCTGTCTGGCGGATTTTCCATTTTAAACCTCTCTCTTAGA	598

## RESULT 7

```

US-10-027-632-41182
; Sequence 41182, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and
; TITLE OF INVENTION: Polymorphisms in

```



446	Qy	ACAGCTAGTGTGAGTCTAATCTTCGCAGAGTCAGGCTCTACCTGTCGTGATGACAAATGGC	505
503	Db	ACAGCTAGTGTGAGTCTAATCTTCGCAGAGTCAGGCTCTACCTGTCGTGATGACAAATGGC	444
506	Qy	ACACTATGTCCTTTAAACCTGATTGAGACACCAAAATGTTTGTCAATATTTTCCCAGGG	565
443	Db	ACACTATGTCCTTTAAACCTGATTGAGACACCAAAATGTTTGTGAATATTTTCCCAGGG	384
566	Qy	AAAAAACCGGAGTAGTCTCTAAATTTCTATATCATCATTAATATAGTTTTTACCTGTGGATT	625
383	Db	AAAAAACCGGAGTAGTCTCTAAATTTCTATATCATCATTAATATAGTTTTTACCTGTGGATT	324
626	Qy	GGAAAAACCCAGCTCTGTATTGCATTTCAAGGCGGGACAGCCTTTGGTGCACTGCTCTGGCG	685
323	Db	GGAAAAACCCAGCTCTGTATTGCATTTCAAGGCGGGACAGCCTTTGGTGCACTGCTCTGGCG	264
686	Qy	GGATTTTCCATTTTAACTCTCTTTAGAAAGCGCCTTCTCATGPHAAAGTTTCTGTATGCCG	745
263	Db	GGATTTTCCATTTTAACTCTCTTTAGAAAGCGCCTTCTCATGPHAAAGTTTCTGTATGCCG	204
746	Qy	CCAGAGCGCCGAGAGAGGGGACGGGGCTGGAGACGCCCGCAGAGGGCTACGTCCCT	805
203	Db	CCAGAGCGCCGAGAGAGGGGACGGGGCTGGAGACGCCCGCAGAGGGCTACGTCCCT	144
806	Qy	GCTGACAGAGTCTCTGCTCTCTCGGGGGGCGAGGCCACTTCCACAACCCCTCGGG	865
143	Db	GCTGACAGAGTCTCTGCTCTCTCGGGGGGCGAGGCCACTTCCACAACCCCTCGGG	84
866	Qy	GAGAACCCCCCAAGGGGAGAGAGGGGCTCGGCCCTTGGCCCCCGAGACCTTCGTCCTCTA	925
83	Db	GAGAACCCCCCAAGGGGAGAGAGGGGCTCGGCCCTTGGCCCCCGAGACCTTCGTCCTCTA	24
926	Qy	GGTCGAGTCTGAATCGGCCCTTG	948
23	Db	GGTCGAGTCTGAATCGGCCCTTG	1

## RESULT 9

US-09-864-761-9446/c  
 ; Sequence 9446, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 ; FILE REFERENCE: Aecomica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30

RESULT 10

US-10-094-749-1376/c  
; Sequence 1376, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOPARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOKYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1376  
; LENGTH: 1904  
; TYPE: DNA  
; ORGANISM: Homo sapiens

Query Match 22.9%; Score 458.4; DB 15; Length 1904;  
Best Local Similarity 99.8%; Pred. No. 2.1e-130;  
Matches 459; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	615	ACCTGTGATTGGAAACCAGCTCTGATTTTCAGTTCAGGGCGGACAGCCTTTGGTGC	674
Dd	460	ATCTGTGATTGGAAACCAGCTCTGATTTTCAGTTCAGGGCGGACAGCCTTTGGTGC	401
Qy	675	ACTGTCTGGCGGATTTTCCATTTTAACCTCTCTAGAGCGCCTTCATGGTAAAGT	734
Dd	400	ACTGTCTGGCGGATTTTCCATTTTAACCTCTCTAGAGCGCCTTCATGGTAAAGT	341
Qy	735	TCTTGATGCCCGCAGGAGCGCCAGGAGAGGCGAGGGGGCTGGAGACGCCCGCAGAGGG	794
Dd	340	TCTTGATGCCCGCAGGAGCGCCAGGAGAGGCGAGGGGGCTGGAGACGCCCGCAGAGGG	281
Qy	795	CTAGTGCCTCTGTGGACAGAGTCTCTGCTCTCTCGGGGGCGGACGCCACCTCCCAC	854
Dd	280	CTAGTGCCTCTGTGGACAGAGTCTCTGCTCTCTCGGGGGCGGACGCCACCTCCCAC	221
Qy	855	AACCCCTCGGGGAGAAGCCCCCAAGGGGAGAGAGCGGGCTGCCCTGCCCGCAGCAC	914
Dd	220	AACCCCTCGGGGAGAAGCCCCCAAGGGGAGAGAGCGGGCTGCCCTGCCCGCAGCAC	161
Qy	915	TTCGCTCTTAGTTCGAGTCTGAATCGGCTTGGAGACCTGTGGCTTCGGGAGCCCC	974
Dd	160	TTCGCTCTTAGTTCGAGTCTGAATCGGCTTGGAGACCTGTGGCTTCGGGAGCCCC	101
Qy	975	TGCAAGAGTCCACAGGCGCGCTCGCTCTCTCTGTTTTATCTCTCCCGACCT	1034
Dd	100	TGCAAGAGTCCACAGGCGCGCTCGCTCTCTCTGTTTTATCTCTCCCGACCT	41
Qy	1035	CTGGCAGGAAACCGCTCATCTGTTAGCGCCCTTTGCGAGCCT	1074
Dd	40	CTGGCAGGAAACCGCTCATCTGTTAGCGCCCTTTGCGAGCCT	1

US-10-094-749-1376  
; TYPE: DNA  
; ORGANISM: Homo sapiens

Query Match 19.2%; Score 383.2; DB 15; Length 508;  
Best Local Similarity 87.8%; Pred. No. 2e-107;  
Matches 455; Conservative 0; Mismatches 53; Indels 10; Gaps 3;

Qy	1472	GAGGGCCCTGGGAACTTCGAGAGCCAGGTCACCTCTCACTCTGTGCTCTTAGTAT	1531
Dd	1	GAGGGCCCTGGGAACTTCGAGAGCCAGGTCACCTCTCACTCTGTGCTCTTAGTAT	60
Qy	1532	CTTGATGCTCTGTGCTTTTTCATAGCTGCTCCCTGACAGGAACTCCATCCCCATCT	1591
Dd	61	CTTGATGCTCTGTGCTTTTTCATAGCTGCTCCCTGACAGGAACTCCATCCCCATCT	120
Qy	1592	TTGCTGCTGTGCGAACCTTCAGAAATCTGCAAGGTCAGCTTAGAGGTCACCTCTCCGG	1651
Dd	121	TTGCTGCTGTGCGAACCTTCAG-AACTGCAAGGTCAGCTTAGAGGTCACCTCTCCGG	179
Qy	1652	AAGCTTTTCTCAACCCCTCCCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1711
Dd	180	AAGCTTTTCTCAAGAGCCCGCCCG--GGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	236
Qy	1712	CAGTATACAGTGTCCGCTCTCCACCTCCACACCTCCACCTCCACCTCCACCTCCAGAGTG	1771
Dd	237	CAGTATACAGTGTCCGCTCTCCACCTCCACACCTCCACCTCCACCTCCACCTCCAGAGTG	296
Qy	1772	AGCCACAGGAGGAGGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1831
Dd	297	AGCCACAGGAGGAGGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	356
Qy	1832	AATGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1891
Dd	357	AATGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	416
Qy	1892	GAGGCTGGATCTGGATCTGCTGCTTTTGTAGTGAAGTGTCCCTTTTAAACAGCAAATGGGCTGGCC	1951
Dd	417	GAGG-----CTGGATCTGCTGCTTTTGTAGTGAAGTGTCCCTTTTAAACAGCAAATGGGCTGGCC	470
Qy	1952	TGGCTCGGGCCCTGCTTTTGTGCTCTCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTG	1989

QY 880 GGGAGGAGACGGGCCCTGGCCCCCTGCCCCGAGCACCTTCGGTCTCTAGGTGGAGTCTGAA 939

১৫

Qy	880	GGGAGGAGCGGCTGGCCCTCCCGAGACACTTCCGTCCTTAGTGGAGCTCGAA	939
Db	139	GGGAGGAGCGGCTGGCCCTCCCGAGACACTTCCGTCCTTAGTGGAGCTCGAA	80
Qy	940	TGGCCCTTGGGACCTGCTTTGCTTCGGGACCCCTGCAAGACTCCACAGGCGCCGTC	999
Db	79	TGGCCCTTGGGACCTGCTTTGCTTCGGGACCCCTGCAAGACTCCACAGGCGCCGTC	20
Qy	1000	GCTCTTCCTCCTGCTTTT	1018
Db	19	GCTCTTCCTCCTGCTTTT	1

```

: APPLICANT: David W. Morris
: APPLICANT: Marc S. Malandro
: TITLE OF INVENTION: Novel Compositions and Methods in Cancer
: FILE REFERENCE: 529452001000
: CURRENT APPLICATION NUMBER: US/10/322,281

```

APPLICANT: David W. Morris  
APPLICANT: Marc S. Malandro

INVENTOR: MAIC S. MALANDRINO  
; AFFILIANT: MAIC S. MALANDRINO  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer

FILE REFERENCE: 529452001000

US/10/322,281

us-09-820-095b-3\_copy\_1\_2000.rnpb

Sun Nov 28 09:38:19 2004

Search completed: November 22, 2004, 08:18:19  
Job time : 991.76 secs

```
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 648
; LENGTH: 51289
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-648

Query Match      5.4%; Score 108.2; DB 17; Length 51289;
Best Local Similarity 79.1%; Pred.No. 1.7e-21;
Matches 148; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

QY 247 GCAACATAGCGAGATAAAATTTAAATTTAGCCAGATGTGGTAGCCCTCTGTAGTCT 306
DB 18038 GCAGACCCCGTCTTACAAAATAAATAAATTTAGCCAGGTGTCCACACACCTGTAGTCC 18097

QY 307 CAGCGACTCAGGAGCTGAGCGAGGAGGCTCACCAGAGTGCA-GAGTTCAAGGATGCAGT 365
DB 18098 CAGCTACTCGGAGGCTGAGCGAGGAGATCACTTTGAGCCCGAGGAGTTCAAGGTTGCAGT 18157

QY 366 GAGCTATGATCCTCCCACTGCACCTGAAAGCTGGGTGACAGCAAGACCCCTGGCTCTAAT 425
DB 18158 GCACTTTGATCATGCCACTGCACCTCCAGCTGGGTGAAAAGCAAGACCCCTATCTCTAAA 18217

QY 426 AATGAATACATAAAGT 442
DB 18218 GAATGCATAATAAATT 18234

RESULT 15
US-09-799-799-3
; Sequence 3, Application US/09799799
; Patent No. US20020132291A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: C0001157
; CURRENT APPLICATION NUMBER: US/09/799,799
; CURRENT FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 88191
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(88191)
; OTHER INFORMATION: n = A,T,C or G
US-09-799-799-3

Query Match      5.2%; Score 103.2; DB 9; Length 88191;
Best Local Similarity 69.8%; Pred.No. 7.8e-20;
Matches 155; Conservative 0; Mismatches 63; Indels 4; Gaps 1;

QY 223 TCAAAAGAAAGCTGGACAACTGGCGCAACATAGCGAGATAAAATTTAAATTTAGCC 282
DB 2226 TCAAGACCAAGCTGGCGCAACATTTGAAATCTGTCTGCAAAAATAACACAAATAAGCC 2285

QY 283 AGAT---GTGGTAGCCCTCTAGTCTCAGCGACTCAGGAGGCTGAGCGAGGAGGCTCA 338
DB 2286 AGTTGTGTGGCTGCGCTGTAGTCCAGCTACTCGGAGACTGAGTGGGAGGATTG 2345

QY 339 CCAGAGTCAGAGTTCAAGGATGAGTGAGTGATGATCTCCCACTGCACCTGAAGCTGG 398
DB 2346 CTTGAACCCAGAGGTTGAGGCTGAGTGAGTGTGATTTGTCCTCACTCTAGCCTGG 2405

QY 399 GTGACAGAGCAAGACCCCTGGCTCTAAATAAATAATACATAAA 440
DB 2406 GTGACAGAGCAAGACCCCTGTCTCAAAAATAAATAATTAATTA 2447
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 166.842 Seconds  
(without alignments)  
8520.488 Million cell updates/sec

Title: US-09-820-095B-3\_COPY\_1\_2000  
Perfect score: 2000  
Sequence: 1 tctccagtcacatgggtgccc.....agctggcgtgcagctgccca 2000

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PTOUT\_COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	159	8.0	394	3	US-09-191-136-27
2	158.4	7.9	160	4	US-08-513-395C-22989
3	95	4.8	59065	3	US-08-813-817-3
4	95	4.8	59065	4	US-09-978-197-3
5	95	4.8	59065	4	US-10-135-696-3
6	92	4.6	1381	2	US-08-454-557C-49
7	92	4.6	1381	2	US-08-340-426D-49
8	92	4.6	1381	2	US-08-450-673C-49
9	92	4.6	1381	5	PCT-US95-17111A-49
10	90.6	4.5	1001	4	US-09-641-638-78
11	90.6	4.5	1001	4	US-10-170-097-78
12	90.6	4.5	35060	3	US-08-814-095-7
13	89.8	4.5	66804	4	US-09-740-041-3
14	89.4	4.5	83450	4	US-08-811-469-3
15	89.4	4.5	83450	4	US-10-370-659-3
16	89	4.5	7676	1	US-08-451-777A-7
17	89	4.5	7676	2	US-08-451-778A-7
18	89	4.5	7676	2	US-08-998-208-7
19	89	4.5	7676	5	PCT-US95-06743-7
20	89	4.5	8133	3	US-09-659-791A-10
21	88.4	4.4	435	4	US-09-621-976-16669
22	88.4	4.4	472	4	US-08-621-976-16668
23	88.4	4.4	1624	2	US-08-852-807-10
24	88.4	4.4	13674	2	US-08-852-807-1
25	88.2	4.4	31000	4	US-09-966-451-10
26	88.2	4.4	64467	4	US-09-803-671B-3
27	88	4.4	18443	3	US-09-078-294-6
28					Sequence 27, Appl
					Sequence 22989, A
					Sequence 3, Appl
					Sequence 3, Appl
					Sequence 3, Appl
					Sequence 49, Appl
					Sequence 49, Appl
					Sequence 49, Appl
					Sequence 49, Appl
					Sequence 78, Appl
					Sequence 78, Appl
					Sequence 7, Appl
					Sequence 3, Appl
					Sequence 3, Appl
					Sequence 3, Appl
					Sequence 7, Appl
					Sequence 7, Appl
					Sequence 7, Appl
					Sequence 7, Appl
					Sequence 10, Appl
					Sequence 16669, A
					Sequence 16668, A
					Sequence 10, Appl
					Sequence 10, Appl
					Sequence 1, Appl
					Sequence 10, Appl
					Sequence 3, Appl
					Sequence 6, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-191-136-27  
; Sequence 27, Application US/09191136B  
; Patent No. 6214581  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Lynch, Kevin J.  
; APPLICANT: Burgard, Edward C.  
; APPLICANT: Van Biesen, T.  
; TITLE OF INVENTION: Nucleic Acids Encoding A Functional  
; TITLE OF INVENTION: Human Purinoreceptor P2X3 and P2X6 And Methods Of Production  
; TITLE OF INVENTION: And Use Thereof  
; FILE REFERENCE: 6293 US.P1  
; CURRENT APPLICATION NUMBER: US/09/191,136B  
; CURRENT FILING DATE: 1998-11-13  
; EARLIER APPLICATION NUMBER: US 09/008,526  
; EARLIER FILING DATE: 1998-01-16  
; EARLIER APPLICATION NUMBER: US 09/008,185  
; EARLIER FILING DATE: 1998-01-16  
; EARLIER APPLICATION NUMBER: US 60/071,298  
; EARLIER FILING DATE: 1998-01-16  
; EARLIER APPLICATION NUMBER: US 60/071,669  
; EARLIER FILING DATE: 1998-01-16  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 27  
; LENGTH: 394  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Universal Amplification Sequencing Primer  
US-09-191-136-27

Query Match 8.0%; Score 159; DB 3; Length 394;  
Best Local Similarity 91.8%; Pred No. 5.4e-33;  
Matches 168; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1818 GGCCAGCAAGGGAATGTAGGAGGGTGGAGGTGCGAGGCGCTGGGATTAGGGGTTG 1877  
Db 1 GGCCAGCGGTCACAGTACGGGGGGGGGGTGCAGGCGAGCTGGGATTAGGGGTTG 60

QY 1878 AGGGCTGGGTGTTGAGGCTGGATCTGCTGTTTAGTGGAGTGTCCCTTTAAACAG 1937  
Db 61 AGGGCTGGGTGTTGAGGCTGGATCTGCTGTTTAGTGGAGTGTCCCTTTAAACAG 120

QY 1938 CAACTGGGCTGCGCTGGCTCGGGCCCTGCTTTCCTCTGCTTTCAGTCGGGTGCGAGTGG 1997  
Db 121 CAACTGGGCTGCGCTGGCTCGGGCCCTGCTTTCCTCTGCTTTCAGTCGGGTGCGAGTGG 180

QY 1998 CCA 2000

Sequence 3, Appl  
Sequence 45, Appl  
Sequence 2, Appl  
Sequence 1, Appl  
Sequence 10, Appl  
Sequence 623, Appl  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 9, Appl  
Sequence 9, Appl  
Sequence 1, Appl  
Sequence 50, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 5, Appl  
Sequence 6, Appl  
Sequence 1, Appl  
Sequence 5, Appl







QY 440 AGTCTCACAGTAGTGGTAG 459  
 Db 605 ATAAAAATTACCAGGTGTAG 586

## RESULT 8

US-08-450-673C-49/c  
 ; Sequence 49, Application US/08450673C  
 ; Patent No. 594888  
 ; GENERAL INFORMATION:  
 ; APPLICANT: de la Monte, Suzanne  
 ; APPLICANT: Wands, Jack R.  
 ; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
 ; TITLE OF INVENTION: of Alzheimer's Disease  
 ; NUMBER OF SEQUENCES: 121  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 ; STREET: 1100 New York Avenue, Suite 600  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20005-3934

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/450,673C  
 FILING DATE: 30-MAY-1995  
 CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
 NAME: Ludwig, Steven R.  
 REGISTRATION NUMBER: 36,203  
 REFERENCE/DOCKET NUMBER: 0609.3840004  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 371-2600  
 TELEFAX: (202) 371-2540  
 INFORMATION FOR SEQ ID NO: 49:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1381 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: both  
 US-08-450-673C-49

Query Match 4.6%; Score 92; DB 2; Length 1381;  
 Best Local Similarity 69.5%; Pred. No. 9.3e-15;  
 Matches 139; Conservative 0; Mismatches 60; Indels 1; Gaps 1;  
 QY 261 TAAAAAATTATTAAATAGCCAGATGTGTAGCCCTCTGTAGTCTCAGCGACTCAGGAG 320  
 Db 785 TAAAAATACAAAAAATTAGTAGGCGTGTGGCGCCTGTAGTCCAGCTACTCAGGAG 726  
 QY 321 GCTGAGCAGGAGGCTCACCAGAGTGC-AGAGTTCAAGGATGCGAGTGTATGATCTCTG 379  
 Db 725 GCTGGGCGCAGGAGATAAATCTGAACCCGGGAGGAGGTTGCGAGTGAAGATTGGG 666  
 QY 380 CCAGTGCACCTGAAAGCTGGGTGACAGAGCAAGACCTGCTCTAATAAATGAATACATAA 439  
 Db 665 CCAGTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATTAATAAATA 606  
 QY 440 AGTCTCACAGTAGTGGTAG 459  
 Db 605 ATAAAAATTACCAGGTGTAG 586

## RESULT 9

PCT-US95-17111A-49/c  
 ; Sequence 49, Application PC/TUS9517111A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: de la Monte, Suzanne  
 ; APPLICANT: Wands, Jack R.

; TITLE OF INVENTION: Neural Thread Protein Gene Expression and  
 ; TITLE OF INVENTION: Detection of Alzheimer's Disease  
 ; NUMBER OF SEQUENCES: 121  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 ; STREET: 1100 New York Avenue, Suite 600  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20005-3934  
 ; COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/17111A  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/340,426  
 FILING DATE: 14-NOV-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ludwig, Steven R.  
 REGISTRATION NUMBER: 36,203  
 REFERENCE/DOCKET NUMBER: 0609.3840002  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 371-2600  
 TELEFAX: (202) 371-2540  
 INFORMATION FOR SEQ ID NO: 49:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1381 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: both  
 PCT-US95-17111A-49

Query Match 4.6%; Score 92; DB 5; Length 1381;  
 Best Local Similarity 69.5%; Pred. No. 9.3e-15;  
 Matches 139; Conservative 0; Mismatches 60; Indels 1; Gaps 1;  
 QY 261 TAAAAAATTATTAAATAGCCAGATGTGTAGCCCTCTGTAGTCTCAGCGACTCAGGAG 320  
 Db 785 TAAAAATACAAAAAATTAGTAGGCGTGTGGCGCCTGTAGTCCAGCTACTCAGGAG 726  
 QY 321 GCTGAGCAGGAGGCTCACCAGAGTGC-AGAGTTCAAGGATGCGAGTGTATGATCTCTG 379  
 Db 725 GCTGGGCGCAGGAGATAAATCTGAACCCGGGAGGAGGTTGCGAGTGAAGATTGGG 666  
 QY 380 CCAGTGCACCTGAAAGCTGGGTGACAGAGCAAGACCTGCTCTAATAAATGAATACATAA 439  
 Db 665 CCAGTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATTAATAAATA 606  
 QY 440 AGTCTCACAGTAGTGGTAG 459  
 Db 605 ATAAAAATTACCAGGTGTAG 586

## RESULT 10

US-09-641-638-78/c  
 ; Sequence 78, Application US/09641638  
 ; Patent No. 6432648  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Blumenfeld, Marta  
 ; APPLICANT: Bougueleret, Lydie  
 ; APPLICANT: Chumakov, Ilya  
 ; APPLICANT: Cohen, Annick  
 ; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
 ; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM  
 ; FILE REFERENCE: GENSET.051CPI  
 ; CURRENT APPLICATION NUMBER: US/09/641,638  
 ; CURRENT FILING DATE: 2000-08-16  
 ; PRIOR APPLICATION NUMBER: US 09/502,330

```
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 09/502,330
; SOFTWARE: Patent.pm
; NUMBER OF SEQ ID NOS: 1304
; SEQ ID NO 78
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 10-224-341 : polymorphic base C or T
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 10-224-341.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 10-224-341.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 161..179
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 561..580
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 10-224-341 potential probe
; NAME/KEY: misc_feature
; LOCATION: 732
; OTHER INFORMATION: n=a, g, c or t
; OTHER INFORMATION: n=a, g, c or t
US-09-641-638-78

Query Match          4.5%; Score 90.6; DB 4; Length 1001;
Best Local Similarity 72.2%; Pred. No. 1.9e-14;
Matches 117; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 287 GTGGTAGCCCTGTAGTCTCAGCGAGCTCAGGAGGCTGAGCGAGGCTCACCAGAGTG 346
Db 748 GTGGTAGCCCTGTANTCCAGTACTCAGGAGCTGAGCGAGGAGTGTGTCAGCC 689

QY 347 CAGAGTTCAAGGATGCGAGTATGATCTGCCACTGCACTGAAAGCTGGGTGACAGA 406
Db 688 CAGGAATTAGGATGCGGTGAGCCATGATTGTGCCACTGCACTGCGAGCTGGTGTG 629

QY 407 GCAGACCTCGCTCTAATAATGATACATTAAGTCTCACA 448
Db 628 GTGAGACCTGTCTCAAAATACAAACAACAACAACA 587

RESULT 11
US-10-170-097-78/c
; Sequence 78, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumentfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
```

```
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 09/502,330
; SOFTWARE: Patent.pm
; NUMBER OF SEQ ID NOS: 1304
; SEQ ID NO 78
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 10-224-341 : polymorphic base C or T
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 10-224-341.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 10-224-341.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 161..179
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 561..580
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 10-224-341 potential probe
; NAME/KEY: misc_feature
; LOCATION: 732
; OTHER INFORMATION: n=a, g, c or t
; OTHER INFORMATION: n=a, g, c or t
US-10-170-097-78

Query Match          4.5%; Score 90.6; DB 4; Length 1001;
Best Local Similarity 72.2%; Pred. No. 1.9e-14;
Matches 117; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 287 GTGGTAGCCCTGTAGTCTCAGCGAGCTCAGGAGGCTGAGCGAGGCTCACCAGAGTG 346
Db 748 GTGGTAGCCCTGTANTCCAGTACTCAGGAGCTGAGCGAGGAGTGTGTCAGCC 689

QY 347 CAGAGTTCAAGGATGCGAGTATGATCTGCCACTGCACTGAAAGCTGGGTGACAGA 406
Db 688 CAGGAATTAGGATGCGGTGAGCCATGATTGTGCCACTGCACTGCGAGCTGGTGTG 629

QY 407 GCAGACCTCGCTCTAATAATGATACATTAAGTCTCACA 448
Db 628 GTGAGACCTGTCTCAAAATACAAACAACAACAACA 587

RESULT 12
US-08-814-095-7
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
```



```
; NAME/KEY: exon
; LOCATION: complement (30816..31011)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30470..30626)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30187..30274)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 14
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29945..30073)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 15
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29664..29856)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 16
; OTHER INFORMATION: /number= 16
US-08-814-095-7

Query Match 4.5%; Score 90.6; DB 3; Length 35060;
Best Local Similarity 77.7%; Pred. No. 9.8e-14;
Matches 122; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Qy 287 GTGTAGCCCTCTAGTCTCAGGACTCAGGAGGCTGAGGAGGCTCACCAGAGTG 346
Db GTGCACACACCTGTAGTCCCGACTCTGCGGAGGCTGAGGAGGCTCAGTGGGCC 8294

Qy 347 CA-GAGTTCAGGATGAGTGTATGATCTCTCCACTGCACTGAAAGCTGGGTGACAG 405
Db CAGGAGTTGAGGCTGAGTGTATGATCTCTCCACTGCACTGAAAGCTGGGTGACAG 8354

Qy 406 AGCAGACCCCTGGCTCTTAATTAATGATACATAAAGT 442
Db AGCAGACCCCTGGCTCTTAATTAATGATACATAAAGT 8391

RESULT 13
US-09-740-041-3/c
; Sequence 3, Application US/09740041
; Patent No. 6562593
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001001
; CURRENT APPLICATION NUMBER: US/09/740,041
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 66804
; TYPE: DNA
; ORGANISM: Human
US-09-740-041-3

Query Match 4.5%; Score 89.8; DB 4; Length 66804;
Best Local Similarity 75.1%; Pred. No. 2.2e-13;
Matches 139; Conservative 0; Mismatches 42; Indels 4; Gaps 2;

Qy 260 ATAAAAATTTTAAATATAGCAGATGTGTAGC---CCCTGTAGTCTCAGGATCA 316
Db ACAGATATTTTAAATATAGCAGATGTGTAGC---CCCTGTAGTCTCAGGATCA 12310

Qy 317 GGAGCTCAGCAGGAGCTCAGGATGCA-GAGTTCAGGATGAGTGTATGAT 375
Db GGAGCTCAGCAGGAGCTCAGGATGCA-GAGTTCAGGATGAGTGTATGAT 12250

; NAME/KEY: exon
; LOCATION: complement (30816..31011)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30470..30626)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30187..30274)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 14
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29945..30073)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 15
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29664..29856)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 16
; OTHER INFORMATION: /number= 16
US-08-814-095-7

Query Match 4.5%; Score 89.4; DB 4; Length 83450;
Best Local Similarity 65.9%; Pred. No. 3.1e-13;
Matches 176; Conservative 0; Mismatches 86; Indels 5; Gaps 3;

Qy 223 TCATAAGAAAGCTGGACACCTGG-GCACATAGGAGATATAAAATTTTAAATTAGC 281
Db TCATAAGAAAGCTGGACACCTGG-GCACATAGGAGATATAAAATTTTAAATTAGC 60760

Qy 282 CAGATGTGTAGC---CCCTGTAGTCTCAGGAGCTCAGGAGGCTGAGCAGGAGCTCA 338
Db CAGATGTGTAGC---CCCTGTAGTCTCAGGAGCTCAGGAGGCTGAGCAGGAGCTCA 60760

Qy 339 CCAGAGTGCAGAGTTCA-AGGATGCGAGTGTATGATCTCCACTGCACTGAAAGCTG 397
Db CCAGAGTGCAGAGTTCA-AGGATGCGAGTGTATGATCTCCACTGCACTGAAAGCTG 60820

Qy 398 GGTGACAGACGAGCCCTGGCTCTAATAAATGAATACATAAAGTCTCACAGCTAGTGT 457
Db GGTGACAGACGAGCCCTGGCTCTAATAAATGAATACATAAAGTCTCACAGCTAGTGT 60880

Qy 458 AGCTAATCTCTGCCAGAGTCAGGCTCT 484
Db AGCTAATCTCTGCCAGAGTCAGGCTCT 60907

RESULT 15
US-10-370-659-3
; Sequence 3, Application US/10370659
; Patent No. 6632949
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001171DIV
; CURRENT APPLICATION NUMBER: US/10/370,659
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 11
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 83450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(83450)
; OTHER INFORMATION: n = A,T,C or G
US-10-370-659-3

Query Match      4.5%; Score 89.4; DB 4; Length 83450;
Best Local Similarity 65.9%; Pred. No. 3.1e-13;
Matches 176; Conservative 0; Mismatches 86; Indels 5; Gaps 3;

Qy 223 TCAAAAGAAAGCTGGACAACCTGG-GCAACATAGCGAGATAAAAAATTATTAAATTAGC 281
Db 60641 TCAAGACCAGCCTGGCCCAACATGGTGAACCTTGTCTCTCTTAAATAACAAAAATTAGC 60700

Qy 282 CAGATGTGGTAGC---CCCCTGTAGTCTCAGCGACTCAGGAGGCTGAGCGAGGCTCA 338
Db 60701 CAGCATGGTGGCCACGACCTGTAGTCCCAGCTACTCAGGAGGCTGAGCGAGGAGATCA 60760

Qy 339 CCAGAGTGCAGAGTTC-AAGGATGAGTGTATGATCCTGCCACTGTGCACTGAAAAGCTG 397
Db 60761 CTTCAACCTGGGAGGCAGAGGTTGCATTGATCAAGATCACGCCACTGCACTGCAGCCTG 60820

Qy 398 GGTGACGAGCAGACCCCTGGCTCTATAAATGAATACATAAAGTCTCAGAGTAGTGGT 457
Db 60821 GGTGATAGACCAAGTCTCCATCTCAAAAAAAAATTTACAAAAGTGGGCCGGTTGTGGT 60880

Qy 458 AGCTAATCTCTGCCAGAGTCAAGGCTCT 484
Db 60881 AGCTATGCCAGTAAATCCAAAGCTCT 60907
```

Search completed: November 21, 2004, 17:43:07  
Job time : 169.842 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 21, 2004, 12:30:45 ; Search time 958.553 Seconds  
(without alignments)  
10952.804 Million cell updates/sec

Title: US-09-820-095B-3\_COPY\_1\_2000  
Perfect score: 2000  
Sequence: 1 tctccaagtcacatgggtgcc.....agctgcggtgcagctgccca 2000

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues  
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2000	100.0	16449	8	AAL51049 Human P2X
2	563	28.1	563	4	AAL18897 Probe #88
3	563	28.1	563	4	ABa63896 Human foe
4	563	28.1	563	4	AAL44018 Probe #12
5	563	28.1	563	4	ABa31076 Probe #95
6	563	28.1	563	4	AAL38120 Human bon
7	563	28.1	563	4	AAL12404 Human bra
8	563	28.1	563	4	ABs37739 Human liv
9	563	28.1	563	6	ABs12139 Human gen
10	554	27.7	554	4	AAL43912 Probe #12
11	554	27.7	554	4	ABa30380 Probe #94
12	554	27.7	554	6	ABs12029 Human gen
13	458.4	22.9	1904	10	ADAs3808 Human cod
14	383.2	19.2	508	12	ACH73486 Human sin
15	257	12.8	268	2	AAH87257 Human bio
16	250.6	12.5	251	2	AAx11014 Human bia
17	159	8.0	394	4	ADa04975 5' end of
18	158.4	7.9	160	3	AAc18914 Human sec
19	139	7.0	139	4	AAI56717 Probe #25
20	139	7.0	139	4	ABa40615 Probe #19
21	139	7.0	139	6	ABs24206 Human gen

ALIGNMENTS

RESULT 1  
AAL51049

ID AAL51049 standard; DNA; 16449 BP.

AC AAL51049;

DT 20-FEB-2003 (first entry)

XX Human P2X-like purinergic receptor G-protein coupled receptor gene.

Human; gene; ds; gene therapy; G-protein coupled receptor; chromosome 22;  
P2X-like purinergic receptor; anaplastic oligodendroglioma; GPCR;  
chronic lymphocytic leukaemia; lung carcinoma; colon carcinoma;  
brain carcinoma; drug screening; tissue typing; pharmacogenomic analysis;  
SNP; single nucleotide polymorphism.

XX Homo sapiens.

Key	Location/Qualifiers
variation	/replace(136,C)
variation	/tag= a
variation	/note= "Single nucleotide polymorphism"
variation	/replace(253,C)
variation	/tag= b
variation	/note= "Single nucleotide polymorphism"
variation	/replace(573,T)
variation	/tag= c
variation	/note= "Single nucleotide polymorphism"
variation	/replace(2000,G)
variation	/tag= d
variation	/note= "Single nucleotide polymorphism"
variation	/replace(2040..13451)
variation	/tag= e
variation	/product= "Human P2X-like purinergic receptor-related G-protein coupled receptor"
variation	/replace(2040..2095)
variation	/tag= f
variation	/number= 1
variation	/tag= g
variation	/number= 1
variation	/replace(2222,C)
variation	/tag= h
variation	/note= "Single nucleotide polymorphism"

22	138.4	6.9	556	12	ACH87204	Ach87204 Human gen	
23	120.6	6.0	121	12	ADK92993	Adk92993 Polynucle	
24	120.6	6.0	121	12	ADK92992	Adk92992 Polynucle	
25	120.6	6.0	121	12	ADK92994	Adk92994 Polynucle	
C	26	109.6	5.5	1356	5	AA878404	AA878404 DNA encod
27	103.2	5.2	88191	8	ABX14763	Abx14763 Genomic D	
28	102.2	5.1	285	5	ABA18022	Abal8022 Human ner	
29	102.2	5.1	285	5	ABA18019	Abal8019 Human ner	
C	30	102	5.1	52616	4	AAK70459	Aak70459 Human imm
C	31	102	5.1	52616	4	AAK78930	Aak78930 Human imm
C	32	102	5.1	304905	11	ADP75180	Adp75180 Human End
C	33	101.8	5.1	2136	4	AAK72074	Aak72074 Human imm
C	34	101.8	5.1	2136	4	AAK72073	Aak72073 Human imm
C	35	101.2	5.1	24908	4	AAK81665	Aak81665 Human imm
C	36	99	5.0	18902	8	ABZ73865	Abz73865 Secreterd
37	99	5.0	18902	10	ABZ67444	Abz67444 Human sec	
38	98.2	4.9	80928	12	ADQ25290	Adq25290 Human pro	
C	39	98	4.9	335	2	AAQ59473	Aaq59473 Human bra
40	97.4	4.9	602	4	AAAL06845	Aal16845 Human rep	
41	97.4	4.9	53000	8	AAAD55325	Aad55325 Human pro	
42	97.4	4.9	139308	8	ADBI2769	Adbi2769 Human PRX	
C	43	97.2	4.9	32190	4	AA836709	Aas36709 Human car
C	44	97.2	4.9	32190	10	ADe47403	Ade47403 Human car
45	96.8	4.8	15870	5	ABA16129	Abal6129 Human ner	

```

FT exon 2777..2927
FT /tag= i
FT /number= 2
FT variation replace(2783,G)
FT /tag= j
FT /note= "Single nucleotide polymorphism"
FT intron 2928..4822
FT /tag= k
FT /number= 2
FT variation replace(3199,A)
FT /tag= l
FT /note= "Single nucleotide polymorphism"
FT variation replace(3307,G)
FT /tag= m
FT /note= "Single nucleotide polymorphism"
FT exon 4823..4894
FT /tag= n
FT /number= 3
FT intron 4895..9510
FT /tag= o
FT /number= 3
FT variation replace(5012,G)
FT /tag= p
FT /note= "Single nucleotide polymorphism"
FT variation replace(6169,C)
FT /tag= q
FT /note= "Single nucleotide polymorphism"
FT variation replace(7647,A)
FT /tag= r
FT /note= "Single nucleotide polymorphism"
FT variation replace(8638,T)
FT /tag= s
FT /note= "Single nucleotide polymorphism"
FT variation replace(9409,T)
FT /tag= t
FT /note= "Single nucleotide polymorphism"
FT exon 9511..9586
FT /tag= u
FT /number= 4
FT intron 9587..9776
FT /tag= v
FT /number= 4
FT exon 9777..9870
FT /tag= w
FT /number= 5
FT intron 9871..9952
FT /tag= x
FT /number= 5
FT exon 9953..10033
FT /tag= y
FT /number= 6
FT intron 10034..10109
FT /tag= z
FT /number= 6
FT exon 10110..10251
FT /tag= aa
FT /number= 7
FT intron 10252..10343
FT /tag= ab
FT /number= 7
FT exon 10344..10453
FT /tag= ac
FT /number= 8
FT intron 10454..12638
FT /tag= ad
FT /number= 8
FT variation replace(10504,C)
FT /tag= ae
FT /note= "Single nucleotide polymorphism"
FT variation replace(12609,A)
FT /tag= af
FT /note= "Single nucleotide polymorphism"
FT exon 12639..12732

```

```

FT /tag= ag
FT /number= 9
FT intron 12733..12844
FT /tag= ah
FT /number= 9
FT exon 12845..12910
FT /tag= ai
FT /number= 10
FT intron 12911..13085
FT /tag= aj
FT /number= 10
FT exon 13086..13163
FT /tag= ak
FT /number= 11
FT intron 13164..13253
FT /tag= al
FT /number= 11
FT exon 13254..13451
FT /tag= am
FT /number= 12
FT variation replace(13367,A)
FT /tag= an
FT /note= "Single nucleotide polymorphism"
FT variation replace(14191,C)
FT /tag= ao
FT /note= "Single nucleotide polymorphism"
FT variation replace(14227,G)
FT /tag= ap
FT /note= "Single nucleotide polymorphism"
FT variation replace(15027,C)
FT /tag= aq
FT /note= "Single nucleotide polymorphism"
FT variation replace(15441,C)
FT /tag= ar
FT /note= "Single nucleotide polymorphism"
FT XX
FT PN WO200279229-A2.
FT XX
FT PD 10-OCT-2002.
FT XX
FT PF 28-MAR-2002; 2002WO-US009545.
FT XX
FT PR 29-MAR-2001; 2001US-00820095.
FT XX
FT PA (PEKE ) PE CORP NY.
FT XX
FT PI Wei M, Gong F, Di Francesco V, Beasley EM;
FT XX
FT XX WPI: 2003-040648/03.
FT DR P-PSDB; AA015997.
FT DR
FT XX
FT PT New peptides related to P2X-like purinergic receptor subfamily, useful for
FT PT treating disorders associated with abnormal expression of protease in
FT PT anaplastic oligodendroglioma, leukemia, carcinoma lung, or large cell
FT PT lung carcinoma.
FT PT
FT PS Claim 4; Fig 3A-E; 87pp; English.
FT XX
FT CC The invention comprises the amino acid and coding sequence (located on
FT CC chromosome 22) of a human G-protein coupled receptor (GPCR) which is
FT CC related to the P2X-like purinergic receptor subfamily. The DNA and
FT CC protein sequences of the invention are useful for treating: anaplastic
FT CC oligodendroglioma; chronic lymphocytic leukaemia; lung carcinoma; colon
FT CC carcinoma; and brain carcinoma. The DNA and protein sequences of the
FT CC invention are also useful for drug screening assays, tissue typing and
FT CC pharmacogenomic analysis. The present genomic DNA sequence represents the
FT CC gene encoding the human GPCR that is related to the P2X-like purinergic
FT CC receptor subfamily
FT XX
FT SQ Sequence 16449 BP; 3284 A; 4727 C; 4589 G; 3849 T; 0 U; 0 Other;

```

Query Match 100.0%; Score 2000; DB 8; Length 16449;  
 Best Local Similarity 100.0%; Pred. No. 0;

		Matches	2000;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;																									
Qy	1	TC	CC	AA	GT	CC	AT	GG	T	GC	CT	GG	T	AG	GA	GA	CA	GG	GG	G	A	T	GA	AT	GT	GA	AA	CC	CT	CA	T	GC	GC	60		
Db	1	TC	CC	AA	GT	CC	AT	GG	T	GC	CT	GG	T	AG	GA	GA	CA	GG	GG	G	A	T	GA	AT	GT	GA	AA	CC	CT	CA	T	GC	GC	60		
Qy	61	T	A	T	A	G	C	C	A	C	T	G	C	T	C	C	C	T	G	C	C	T	G	C	A	T	C	A	C	T	G	C	C	T	120	
Db	61	T	A	T	A	G	C	C	A	C	T	G	C	T	C	C	C	T	G	C	C	T	G	C	A	T	C	A	C	T	G	C	C	T	120	
Qy	121	A	A	G	C	A	C	T	G	C	T	T	C	T	T	A	G	G	A	C	A	C	T	G	C	C	G	C	A	T	A	T	G	A	A	180
Db	121	A	A	G	C	A	C	T	G	C	T	T	C	T	T	A	G	G	A	C	A	C	T	G	C	C	G	C	A	T	A	T	G	A	A	180
Qy	181	G	C	T	A	A	G	C	A	C	G	A	A	A	T	C	T	T	T	C	T	T	A	A	G	T	C	A	T	A	C	A	G	T	G	240
Db	181	G	C	T	A	A	G	C	A	C	G	A	A	A	T	C	T	T	T	C	T	T	A	A	G	T	C	A	T	A	C	A	G	T	G	240
Qy	241	A	C	T	G	G	C	A	A	T	A	T	T	T	T	A	A	A	T	T	T	A	A	T	T	A	A	T	T	A	T	T	A	T	300	
Db	241	A	C	T	G	G	C	A	A	T	A	T	T	T	T	A	A	A	T	T	T	A	A	T	T	A	A	T	T	A	T	T	A	T	300	
Qy	301	T	A	G	T	C	T	C	A	G	A	C	T	C	A	G	A	G	G	T	C	A	C	A	G	A	G	T	C	A	G	A	G	T	360	
Db	301	T	A	G	T	C	T	C	A	G	A	C	T	C	A	G	A	G	G	T	C	A	C	A	G	A	G	T	C	A	G	A	G	T	360	
Qy	361	G	C	A	G	T	A	G	C	T	A	T	C	C	G	C	A	T	G	C	A	T	G	A	A	G	T	G	G	T	C	A	G	A	C	420
Db	361	G	C	A	G	T	A	G	C	T	A	T	C	C	G	C	A	T	G	C	A	T	G	A	A	G	T	G	G	T	C	A	G	A	C	420
Qy	421	C	T	A	T	A	A	A	T	G	A	T	C	T	C	A	C	A	G	T	C	A	C	A	G	T	A	A	T	C	C	T	G	C	480	
Db	421	C	T	A	T	A	A	A	T	G	A	T	C	T	C	A	C	A	G	T	C	A	C	A	G	T	A	A	T	C	C	T	G	C	480	
Qy	481	C	T	C	A	C	C	T	G	T	A	G	A	C	A	A	T	G	C	A	C	T	A	T	T	T	A	A	C	T	G	A	T	G	540	
Db	481	C	T	C	A	C	C	T	G	T	A	G	A	C	A	A	T	G	C	A	C	T	A	T	T	T	A	A	C	T	G	A	T	G	540	
Qy	541	T	G	T	T	T	G	T	G	A	T	A	T	T	T	C	C	C	A	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	600	
Db	541	T	G	T	T	T	G	T	G	A	T	A	T	T	T	C	C	C	A	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	600	
Qy	601	A	T	T	A	T	A	T	A	T	T	A	C	T	G																					

Qy	1081	CTGAGCGGAGACCGCTTGGCGCCTCACTTAGAGCGGACCCCGGGATGTGGCGGAGTC	1141
Db	1081	CTGAGCGGAGACCGCTTGGCGCCTCACTTAGAGCGGACCCCGGGATGTGGCGGAGTC	1140
Qy	1141	TGGGGCTCGCCTGACCAATCGAGGTGTGGCGTCCATCGACTGGCGTCTGCCACGGCAATT	1200
Db	1141	TGGGGCTCGCCTGACCAATCGAGGTGTGGCGTCCATCGACTGGCGTCTGCCACGGCAATT	1200
Qy	1201	GCACGCGCTCCCGCGGGCGGTGCGCCCGGCAACCCAGTGTGTAGTTTGGCGTGA	1260
Db	1201	GCACGCGCTCCCGCGGGCGGTGCGCCCGGCAACCCAGTGTGTAGTTTGGCGTGA	1260
Qy	1261	CCGTGGCTCTCTGCGCTGAGGCTCTCGCTCAGAGGATAAATCGACGCGCCACGGGC	1320
Db	1261	CCGTGGCTCTCTGCGCTGAGGCTCTCGCTCAGAGGATAAATCGACGCGCCACGGGC	1320
Qy	1321	TATGCACTGGGCTGGCGCCTTGTGGGCATCTCTCGCTGCTTCTTAGGGGGTTCCAGCAT	1380
Db	1321	TATGCACTGGGCTGGCGCCTTGTGGGCATCTCTCGCTGCTTCTTAGGGGGTTCCAGCAT	1380
Qy	1381	CGCCCCCTTTCGTGAGCTGGGAACACGCGCTGACTCCAGGACTCTGTGTGTCCTCACATG	1440
Db	1381	CGCCCCCTTTCGTGAGCTGGGAACACGCGCTGACTCCAGGACTCTGTGTGTCCTCACATG	1440
Qy	1441	CACCTGGGGAAGTGGCGGGGGCAGCTTTTCAGAGGGGCTTGGGGAATTCGCAGAGCCAG	1500
Db	1441	CACCTGGGGAAGTGGCGGGGGCAGCTTTTCAGAGGGGCTTGGGGAATTCGCAGAGCCAG	1500
Qy	1501	GTCAACCTCTCACTCTGTGGCTCTTAGTTATCTTGTCATGCTGTGGTCTTTGTGATACGCTG	1560
Db	1501	GTCAACCTCTCACTCTGTGGCTCTTAGTTATCTTGTCATGCTGTGGTCTTTGTGATACGCTG	1560
Qy	1561	CTCCCTGCACGAGGAACCTCCATCCCACTTTTGTCTGTTGTGGAATTCAGAAATCTG	1620
Db	1561	CTCCCTGCACGAGGAACCTCCATCCCACTTTTGTCTGTTGTGGAATTCAGAAATCTG	1620
Qy	1621	CAAGGCTCAGCTTAGAGTCACTTCTTCGGGAAGCTTTCCTCAACACCTCCCGGCCCTG	1680
Db	1621	CAAGGCTCAGCTTAGAGTCACTTCTTCGGGAAGCTTTCCTCAACACCTCCCGGCCCTG	1680
Qy	1681	CTGCTGTCGCCCTCAGGCCCTCTCTCACAGCACTGATAACAGCTGTCCGTCTCCACCCT	1740
Db	1681	CTGCTGTCGCCCTCAGGCCCTCTCTCACAGCACTGATAACAGCTGTCCGTCTCCACCCT	1740
Qy	1741	CCACCACTTCACTCCCAACCCAGGAAGTGAAGCCAGAGGGCAGGGAAGAGCTCTCTGC	1800
Db	1741	CCACCACTTCACTCCCAACCCAGGAAGTGAAGCCAGAGGGCAGGGAAGAGCTCTCTGC	1800
Qy	1801	TGTTCTGTGTGCCAGGGCCAGCAAGGAATGTAGGAGGTTGGGAGGTGCAAGGCA	1860
Db	1801	TGTTCTGTGTGCCAGGGCCAGCAAGGAATGTAGGAGGTTGGGAGGTGCAAGGCA	1860
Qy	1861	GCTGGGATTAGGGTTGAGGGCTGGGTGTGTGAGGCTTGGATCTGGATCTCTGCTTTAGTGG	1920
Db	1861	GCTGGGATTAGGGTTGAGGGCTGGGTGTGTGAGGCTTGGATCTGGATCTCTGCTTTAGTGG	1920
Qy	1921	AAGTGTCCCTTTAAAGCAACTGGCCTGGCCTGGCTCGGGCCCTGCTTTGCTCTCTGTTTC	1980
Db	1921	AAGTGTCCCTTTAAAGCAACTGGCCTGGCCTGGCTCGGGCCCTGCTTTGCTCTCTGTTTC	1980
Qy	1981	AGCTGGGCTGCAGCTGCCA	2000
Db	1981	AGCTGGGCTGCAGCTGCCA	2000

RESULT 2	
AAI1897/c	
ID	AAI1897 standard; DNA; 563 BP.
XX	
AC	AAI1897;
XX	
DT	12-OCT-2001 (first entry)

XX DE Probe #8830 for gene expression analysis in human cervical cell sample.  
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 XX KW cervical cancer; ss.  
 XX OS Homo sapiens.  
 XX PN WO200157278-A2.  
 XX PD 09-AUG-2001.  
 XX PF 30-JAN-2001; 2001WO-US000670.  
 XX PR 04-FEB-2000; 2000US-0180312P.  
 XX PR 26-MAY-2000; 2000US-0207456P.  
 XX PR 30-JUN-2000; 2000US-00608408.  
 XX PR 03-AUG-2000; 2000US-00632366.  
 XX PR 21-SEP-2000; 2000US-0234687P.  
 XX PR 27-SEP-2000; 2000US-0236359P.  
 XX PR 04-OCT-2000; 2000GB-00024263.  
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX DR WPI; 2001-489901/53.  
 XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 XX PT gene expression in human cervical epithelial cells.  
 XX PS Claim 25; SEQ ID NO 8830; 487pp; English.  
 XX CC The present invention relates to human single exon nucleic acid probes  
 CC (SNP). The present sequence is one such probe. The SNPs are derived  
 CC from human HeLa cells. The SNPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging of  
 CC diseases of the cervix, notably cervical cancer. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 563 BP; 129 A; 148 C; 156 G; 130 T; 0 U; 0 Other;  
 Query Match 28.1%; Score 563; DB 4; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-139;  
 Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 386 CACTGAAAGCTGGGTGACAGAGCAAGACCTGGCTCTAATAATGAATACATAAAGTCTC 445  
 Db 563 CACTGAAAGCTGGGTGACAGAGCAAGACCTGGCTCTAATAATGAATACATAAAGTCTC 504  
 QY 446 ACAGCTAGTGTAGCTAATCCTGCCAGAGTCAGGCTCTACCTGTCTGATGACAAATGGC 505  
 Db 503 ACAGCTAGTGTAGCTAATCCTGCCAGAGTCAGGCTCTACCTGTCTGATGACAAATGGC 444  
 QY 506 ACACATGTCTTTTAACTGATTCAGACCAACCAATGTTTGTGAATATTTTCCACAGG 565  
 Db 443 ACACATGTCTTTTAACTGATTCAGACCAACCAATGTTTGTGAATATTTTCCACAGG 384  
 QY 566 AAAAAACCGGAAGTAGTCTTAAATCTATACATCCATTATATTAGTTTACCTGTGGATT 625  
 Db 383 AAAAAACCGGAAGTAGTCTTAAATCTATACATCCATTATATTAGTTTACCTGTGGATT 324  
 QY 626 GGGAAAAACCCAGCTCTGATTCGATTCAGGCGGGAGCAGCTTTTGGTGCACTGTCTGGCG 685  
 Db 323 GGGAAAAACCCAGCTCTGATTCGATTCAGGCGGGAGCAGCTTTTGGTGCACTGTCTGGCG 264  
 QY 686 GGATTTTCCATTTTAACTCTTACAGGCGCTTCTCATGCTTAAAGTTCCCTGATGCGCG 745  
 Db 263 GGATTTTCCATTTTAACTCTTACAGGCGCTTCTCATGCTTAAAGTTCCCTGATGCGCG 204

QY 746 CCAGAGCGCCGAGAGAGAGGGGCTGGAGAGCGCCCGCAGAGGGCTACCTGCCT 805  
 Db 203 CCAGAGCGCCGAGAGAGAGGGGCTGGAGAGCGCCCGCAGAGGGCTACCTGCCT 144  
 QY 806 GCTGACAGAGGTCTCTCTGCTCTCTCGCGCGGCGGAGCCACCTCCACACACCCCTGCGG 865  
 Db 143 GCTGACAGAGGTCTCTCTGCTCTCTCGCGCGGCGGAGCCACCTCCACACACCCCTGCGG 84  
 QY 866 GAGAAAGCCCCCAAGGGGAGAGAGCGGCTGCGCCCTGCGCCCGGAGCAGCCTTCCCTCTCTA 925  
 Db 83 GAGAAAGCCCCCAAGGGGAGAGAGCGGCTGCGCCCTGCGCCCGGAGCAGCCTTCCCTCTCTA 24  
 QY 926 GGTGGAGTCTGAATCGGCCTTG 948  
 Db 23 GGTGGAGTCTGAATCGGCCTTG 1

## RESULT 3

ABA63896/c

ID ABA63896 standard; DNA; 563 BP.

XX ABA63896;

AC ABA63896;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #12201.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human foetal liver.

XX Claim 1; SEQ ID NO 12201; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring

XX human gene expression in a sample derived from human foetal liver. The

XX single exon nucleic acid probes may be used for predicting, measuring and

XX displaying gene expression in samples derived from human foetal liver. The

XX present sequence is a single exon nucleic acid probe of the invention.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 563 BP; 129 A; 148 C; 156 G; 130 T; 0 U; 0 Other;

Query Match 28.1%; Score 563; DB 4; Length 563;

Best Local Similarity 100.0%; Pred. No. 1.7e-139;

Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 CACTGAAAGCTGGGTGACAGAGCAAGACCTGGCTCTAATAATGAATACATAAAGTCTC 445

Db 563 CACTGAAAGCTGGTGACAGAGCAAGCCCTGGCTCTAATAAATGAATACATAAAGTCTC 504  
Qy 446 ACAGCTAGTGGTAAATCCCTCCAGAGTCAAGGCTCTACCTGCTGATGACAAATGCG 505  
Db 503 ACAGCTAGTGGTAAATCCCTCCAGAGTCAAGGCTCTACCTGCTGATGACAAATGCG 444  
Qy 506 ACAGCTAGTGGTAAATCCCTCCAGAGTCAAGGCTCTACCTGCTGATGACAAATGCG 565  
Db 443 ACAGCTAGTGGTAAATCCCTCCAGAGTCAAGGCTCTACCTGCTGATGACAAATGCG 384  
Qy 566 AAAAAACCGGAGTGTCTTAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCT 625  
Db 383 AAAAAACCGGAGTGTCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCT 324  
Qy 626 GGGAAAAACCGAGCTCTGATTTGCAATTTACGGGGGAGACAGCCCTTGGTGCATGCTGGCG 685  
Db 323 GGGAAAAACCGAGCTCTGATTTGCAATTTACGGGGGAGACAGCCCTTGGTGCATGCTGGCG 264  
Qy 686 GGATTTTCCATTTAAACCTCTCTTCTAGAAAGCCCTTCTCATGGTAAAGTTCCTGATGCGG 745  
Db 263 GGATTTTCCATTTAAACCTCTCTTCTAGAAAGCCCTTCTCATGGTAAAGTTCCTGATGCGG 204  
Qy 746 CCAGAGAGCCCGAGAGAGGGGAGGGGCTGGAGAGCCCGCCAGAGGGCTACGTGCCCT 805  
Db 203 CCAGAGAGCCCGAGAGAGGGGAGGGGCTGGAGAGCCCGCCAGAGGGCTACGTGCCCT 144  
Qy 806 GCTGACAGAGTCTCTGCTCCCTCTGCGGGGCGCCAGAGCCCGCCAGAGGGCTACGTGCCCT 865  
Db 143 GCTGACAGAGTCTCTGCTCCCTCTGCGGGGCGCCAGAGCCCGCCAGAGGGCTACGTGCCCT 84  
Qy 866 GAGAGAGCCCGAGAGAGGGGAGGGGCTGGAGAGCCCGCCAGAGGGCTACGTGCCCT 925  
Db 83 GAGAGAGCCCGAGAGAGGGGAGGGGCTGGAGAGCCCGCCAGAGGGCTACGTGCCCT 24  
Qy 926 GGTGAGTCTGAATCGGCCCTTG 948  
Db 23 GGTGAGTCTGAATCGGCCCTTG 1

## RESULT 4

AAI44018/c  
ID AAI44018 standard; DNA; 563 BP.  
XX  
AC AAI44018;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Probe #12704 used to measure gene expression in human placenta sample.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US0000663.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX

DR WPI; 2001-488897/53.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human placenta.  
XX  
PS Claim 25; SEQ ID NO 12704; 654pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENP).  
CC The present sequence is one such probe. The probes are useful for  
CC producing a microarray for predicting, measuring and displaying gene  
CC expression in samples derived from human placenta. The probes are useful  
CC for antenatal diagnosis of human genetic disorders  
XX  
SQ Sequence 563 BP; 129 A; 148 C; 156 G; 130 T; 0 U; 0 Other;

Query Match 28.1%; Score 563; DB 4; Length 563;  
Best Local Similarity 100.0%; Pred. No. 1.7e-139;  
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 386 CACTGAAAGCTGGTGTACAGAGCAAGCCCTGGCTCTAATAAATGAATACATAAAGTCTC 445  
Db 563 CACTGAAAGCTGGTGTACAGAGCAAGCCCTGGCTCTAATAAATGAATACATAAAGTCTC 504  
Qy 446 ACAGCTAGTGGTAAATCCCTCCAGAGTCAAGGCTCTACCTGCTGATGACAAATGCG 505  
Db 503 ACAGCTAGTGGTAAATCCCTCCAGAGTCAAGGCTCTACCTGCTGATGACAAATGCG 444  
Qy 506 ACAGCTAGTGGTAAATCCCTCCAGAGTCAAGGCTCTACCTGCTGATGACAAATGCG 565  
Db 443 ACAGCTAGTGGTAAATCCCTCCAGAGTCAAGGCTCTACCTGCTGATGACAAATGCG 384  
Qy 566 AAAAAACCGGAGTGTCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCT 625  
Db 383 AAAAAACCGGAGTGTCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCTTAAATCT 324  
Qy 626 GGGAAAAACCGAGCTCTGATTTGCAATTTACGGGGGAGACAGCCCTTGGTGCATGCTGGCG 685  
Db 323 GGGAAAAACCGAGCTCTGATTTGCAATTTACGGGGGAGACAGCCCTTGGTGCATGCTGGCG 264  
Qy 686 GGATTTTCCATTTAAACCTCTCTTCTAGAAAGCCCTTCTCATGGTAAAGTTCCTGATGCGG 745  
Db 263 GGATTTTCCATTTAAACCTCTCTTCTAGAAAGCCCTTCTCATGGTAAAGTTCCTGATGCGG 204  
Qy 746 CCAGAGAGCCCGAGAGAGGGGAGGGGCTGGAGAGCCCGCCAGAGGGCTACGTGCCCT 805  
Db 203 CCAGAGAGCCCGAGAGAGGGGAGGGGCTGGAGAGCCCGCCAGAGGGCTACGTGCCCT 144  
Qy 806 GCTGACAGAGTCTCTGCTCCCTCTGCGGGGCGCCAGAGCCCGCCAGAGGGCTACGTGCCCT 865  
Db 143 GCTGACAGAGTCTCTGCTCCCTCTGCGGGGCGCCAGAGCCCGCCAGAGGGCTACGTGCCCT 84  
Qy 866 GAGAGAGCCCGAGAGAGGGGAGGGGCTGGAGAGCCCGCCAGAGGGCTACGTGCCCT 925  
Db 83 GAGAGAGCCCGAGAGAGGGGAGGGGCTGGAGAGCCCGCCAGAGGGCTACGTGCCCT 24  
Qy 926 GGTGAGTCTGAATCGGCCCTTG 948  
Db 23 GGTGAGTCTGAATCGGCCCTTG 1

## RESULT 5

ABA31076/c  
ID ABA31076 standard; DNA; 563 BP.  
XX  
AC ABA31076;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Probe #9542 for gene expression analysis in human heart cell sample.  
XX  
KW Human; gene expression; heart; microarray; vascular system; probe;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease; ss.



Db 503 ACAGCTAGTGTAGTAATCTCTGACAGAGTCAGGCTCTACCTGCTCTGATGACAAATGGC 444  
 Qy 506 ACACATATGCTCTTTAACTGATTCAGACACCAAAATGTTTGTGAATTTTCCCGAGG 565  
 Db 443 ACACATATGCTCTTTAACTGATTCAGACACCAAAATGTTTGTGAATTTTCCCGAGG 394  
 Qy 566 AAAAAACCGGAGTAGTCTTAAATCTATACATCCATATATAGTTTACCTGTGGATT 625  
 Db 393 AAAAAACCGGAGTAGTCTTAAATCTATACATCCATATATAGTTTACCTGTGGATT 324  
 Qy 626 GGGAAAAACCGAGCTCTGATTCGATTCAGGGGGGACAGCCCTTGTGTCACCTGTCTGGCG 685  
 Db 323 GGGAAAAACCGAGCTCTGATTCGATTCAGGGGGGACAGCCCTTGTGTCACCTGTCTGGCG 264  
 Qy 686 GGATTTTTCATTTAACTCTCTCTAGAGGCGCTTCTCATGGTAAAGTCTCTGATGCG 745  
 Db 263 GGATTTTTCATTTAACTCTCTCTAGAGGCGCTTCTCATGGTAAAGTCTCTGATGCG 204  
 Qy 746 CCAGAGCGCCGAGGAGGCGGAGCGGCTGAGACGCGCCCGCAGAGGGCTACGTGCCCT 805  
 Db 203 CCAGAGCGCCGAGGAGGCGGAGCGGCTGAGACGCGCCCGCAGAGGGCTACGTGCCCT 144  
 Qy 806 GGTGACACAGAGTCTCTCTGCTCTCTGCGGGGCGGCGGCGGCGGCGGCGGCGGCGG 865  
 Db 143 GGTGACACAGAGTCTCTCTGCTCTCTGCGGGGCGGCGGCGGCGGCGGCGGCGGCGG 84  
 Qy 866 GAGAAAGCCCCAAGGGGAGGAGCGGCGCTGCGCCCTGCGCGGAGCACCTTCCGTCTCTA 925  
 Db 83 GAGAAAGCCCCAAGGGGAGGAGCGGCGCTGCGCCCTGCGCGGAGCACCTTCCGTCTCTA 24  
 Qy 926 GGTGAGTCTGAATCGGCTTG 948  
 Db 23 GGTGAGTCTGAATCGGCTTG 1

## RESULT 7

AAK12404/C  
 ID AAK12404 standard; DNA; 563 BP.  
 AC AAK12404;  
 XX  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe SEQ ID NO: 12395.  
 XX  
 KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
 KW ss.  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483446/52.

Single exon nucleic acid probes for analyzing gene expression in human  
 brains.

XX

Example 4; SEQ ID NO 12395; 650pp + Sequence Listing; English.

PS  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is one of the probes of the  
 CC invention

SQ Sequence 563 BP; 129 A; 148 C; 156 G; 130 T; 0 U; 0 Other;

## Query Match

28.1%; Score 563; DB 4; Length 563;

Best Local Similarity 100.0%; Pred. No. 1.7e-139;

Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 386 CACTGAAAGCTGGTGACAGAGCAAGACCCCTGCTCTAATAATGAATACATAAAGTCTC 445  
 Db 563 CACTGAAAGCTGGTGACAGAGCAAGACCCCTGCTCTAATAATGAATACATAAAGTCTC 504  
 Qy 446 ACAGCTAGTGGTAGCTAATCTGCCAGAGTCAGGCTCTACCTGTCTGATGACAAATGGC 505  
 Db 503 ACAGCTAGTGGTAGCTAATCTGCCAGAGTCAGGCTCTACCTGTCTGATGACAAATGGC 444  
 Qy 506 ACATATGCTCTTTTAACTGATTCAGACCAAAATGTTTGTGAATATTTTCCCGAGG 565  
 Db 443 ACATATGCTCTTTTAACTGATTCAGACCAAAATGTTTGTGAATATTTTCCCGAGG 394  
 Qy 566 AAAAAACCGGAGTAGTCTTAAATCTATACATCCATATATAGTTTACCTGTGGATT 625  
 Db 393 AAAAAACCGGAGTAGTCTTAAATCTATACATCCATATATAGTTTACCTGTGGATT 324  
 Qy 626 GGGAAAAACCGAGCTCTGATTCGATTCAGGGGGGACAGCCCTTGTGCACTGTCTGGCG 685  
 Db 323 GGGAAAAACCGAGCTCTGATTCGATTCAGGGGGGACAGCCCTTGTGCACTGTCTGGCG 264  
 Qy 686 GGATTTTTCATTTAACTCTCTCTAGAGGCGCTTCTCATGGTAAAGTCTCTGATGCGG 745  
 Db 263 GGATTTTTCATTTAACTCTCTCTAGAGGCGCTTCTCATGGTAAAGTCTCTGATGCGG 204  
 Qy 746 CCAGGAGCGCCGAGGAGGCGGCGGCTGAGACGCGCCCGCAGAGGGCTACGTGCCCT 805  
 Db 203 CCAGGAGCGCCGAGGAGGCGGCGGCTGAGACGCGCCCGCAGAGGGCTACGTGCCCT 144  
 Qy 806 GCTGGAACAGAGTCTCTCTGCTCTCTGCGGGGCGGCGGCGGCGGCGGCGGCGGCGG 865  
 Db 143 GCTGGAACAGAGTCTCTCTGCTCTCTGCGGGGCGGCGGCGGCGGCGGCGGCGGCGG 84  
 Qy 866 GAGAAAGCCCCAAGGGGAGGAGCGGCGCTGCGCCCTGCGCGGAGCACCTTCCGTCTCTA 925  
 Db 83 GAGAAAGCCCCAAGGGGAGGAGCGGCGCTGCGCCCTGCGCGGAGCACCTTCCGTCTCTA 24  
 Qy 926 GGTGAGTCTGAATCGGCTTG 948  
 Db 23 GGTGAGTCTGAATCGGCTTG 1

## RESULT 8

ABS37739/c

ID ABS37739 standard; DNA; 563 BP.

XX ABS37739;

AC ABS37739;

XX 25-FEB-2003 (first entry)

XX Human liver single exon probe, SEQ ID No 12729.

XX Human; single exon nucleic acid probe; liver; cirrhosis;

XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;

XX coronary heart disease; ss.

XX Homo sapiens.

OS





CC array; identifying exons in a eukaryotic genome, comprising (a)  
CC algorithmically predicting at least one exon from genomic sequences of  
CC the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene expression  
CC analysis, and for identifying exons in a gene, particularly using human  
CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
CC Karagenes syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
CC present sequence is a single exon probe of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 563 BP; 129 A; 148 C; 156 G; 130 T; 0 U; 0 Other;

Query Match 28.1%; Score 563; DB 6; Length 563;  
Best Local Similarity 100.0%; Pred. No. 1.7e-139;  
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 386 CACTCAAGCTGGGTGACAGACAGACCCCTGGCTCTAATAAATGAATACATAAGTCTC 445  
DB 563 CACTCAAGCTGGGTGACAGACAGACCCCTGGCTCTAATAAATGAATACATAAGTCTC 504  
QY 446 ACAGCTAGTGGTAGTAATCTCTGCGAGAGTCAAGGCTCTACTGCTGATGACAAATGGC 505  
DB 503 ACAGCTAGTGGTAGTAATCTCTGCGAGAGTCAAGGCTCTACTGCTGATGACAAATGGC 444  
QY 506 ACACATATCTCTTTAACTGATGTCAGACACCAATGTTTGTGAATTTTCCCGAGG 565  
DB 443 ACACATATCTCTTTAACTGATGTCAGACACCAATGTTTGTGAATTTTCCCGAGG 384  
QY 566 AAAAAACCGGAGTAGTCTAAATCTATACATCCATTATATTAGTTTACCTCTGGATT 625  
DB 383 AAAAAACCGGAGTAGTCTAAATCTATACATCCATTATATTAGTTTACCTCTGGATT 324  
QY 626 GGGAAACCCAGCTCTGATTGCAATTCAGGGGGGAGACCTTTGGTCACTGTCTGGCG 685  
DB 323 GGGAAACCCAGCTCTGATTGCAATTCAGGGGGGAGACCTTTGGTCACTGTCTGGCG 264  
QY 686 GGATTTTCCATTTTAACTCTCTCTTAGAAGGCGCTTCTCATGCTAAAGTCTCTGATGCG 745  
DB 263 GGATTTTCCATTTTAACTCTCTCTTAGAAGGCGCTTCTCATGCTAAAGTCTCTGATGCG 204  
QY 746 CCAGAGGCGCCGAGAGAGGGGAGGGGCTGAGACGCCCGCCGAGAGGGTACGTGCCT 805  
DB 203 CCAGAGGCGCCGAGAGAGGGGAGGGGCTGAGACGCCCGCCGAGAGGGTACGTGCCT 144  
QY 806 GCTGACAGAGGTCTCTGCTCTCTGCGGGCGGCGAGCCACCTCCCAACACCCCTGGG 865  
DB 143 GCTGACAGAGGTCTCTGCTCTCTGCGGGCGGCGAGCCACCTCCCAACACCCCTGGG 84  
QY 866 GAGAGCCCCCAAGGGGAGAGACGGGCTGGCCCTGGCCCTGGCCGAGACCTTCCGCTCTA 925  
DB 83 GAGAGCCCCCAAGGGGAGAGACGGGCTGGCCCTGGCCCTGGCCGAGACCTTCCGCTCTA 24  
QY 926 GGTGAGGTCTGAATCGGCTTG 948  
DB 23 GGTGAGGTCTGAATCGGCTTG 1

## RESULT 10

AAI43912/c  
ID AAI43912 standard; DNA; 554 BP.  
XX  
XX AAI43912;  
AC  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Probe #12598 used to measure gene expression in human placenta sample.  
XX  
XX Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder; ss.  
XX  
XX Homo sapiens.  
OS  
PN WO200157272-A2.  
XX  
XX 09-AUG-2001.  
PD  
XX 30-JAN-2001; 2001WO-US0000663.  
PF  
XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488897/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human placenta.  
XX  
XX Claim 25; SEQ ID NO 12598; 654pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes (SENP).  
CC The present sequence is one such probe. The probes are useful for  
CC producing a microarray for predicting, measuring and displaying gene  
CC expression in samples derived from human placenta. The probes are useful  
CC for antenatal diagnosis of human genetic disorders  
XX  
XX Sequence 554 BP; 109 A; 170 C; 186 G; 89 T; 0 U; 0 Other;  
SQ  
Query Match 27.7%; Score 554; DB 4; Length 554;  
Best Local Similarity 100.0%; Pred. No. 4.2e-137;  
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 615 ACCTGTGGATTGGGAAACCCAGCTCTGATTGCAATTTTAACTCTCTAGAGCGCTTCTCATGTAAGT 734  
DB 554 ACCTGTGGATTGGGAAACCCAGCTCTGATTGCAATTTTAACTCTCTAGAGCGCTTCTCATGTAAGT 435  
QY 675 ACTGTCTGGCGGATTTTCCATTTTAACTCTCTAGAGCGCTTCTCATGTAAGT 734  
DB 494 ACTGTCTGGCGGATTTTCCATTTTAACTCTCTAGAGCGCTTCTCATGTAAGT 435  
QY 735 TCCTGTATCCCGCCAGAGGCGCCGAGGAGAGGGGCTGGAGACGCCCGCCAGAGGG 794  
DB 434 TCCTGTATCCCGCCAGAGGCGCCGAGGAGAGGGGCTGGAGACGCCCGCCAGAGGG 375  
QY 795 CTACGTGCTCTGTGACACAGGTCTCTGCTCTCTCTGCGGGCGCGAGCCACCTCCAC 854  
DB 374 CTACGTGCTCTGTGACACAGGTCTCTGCTCTCTCTGCGGGCGCGAGCCACCTCCAC 315  
QY 855 AACCCCTGCGGAGAGACCCCGCCAGGAGGAGACCGGCTGCGCCCTGCGCGAGCACC 914  
DB 314 AACCCCTGCGGAGAGACCCCGCCAGGAGGAGACCGGCTGCGCCCTGCGCGAGCACC 255



Db	494	ACTGTCTGGCGGATTTTCATTTTAACTCTCTTAGAAGCGCCTTCTCATGGTAAAGT	435
Qy	735	TCCTGATGCCCGCAGGAGCGCCGAGGAGAGGGGAGGGGGCTGGAGAGCGCCCGCAGAGGG	794
Db	434	TCCTGATGCCCGCAGGAGCGCCGAGGAGAGGGGAGGGGGCTGGAGAGCGCCCGCAGAGGG	375
Qy	795	CTACGTGCCCTGCTGAGACAGAGTCTCTGCTCTCTCGGGCGGCGCAGCCACTCTCCAC	854
Db	374	CTACGTGCCCTGCTGAGACAGAGTCTCTGCTCTCTCGGGCGGCGCAGCCACTCTCCAC	315
Qy	855	AACCCCTGCGGGAGAGACCCCAAGGGAGGAGACGGGCTGGCCCTTGCCCGGAGCACC	914
Db	314	AACCCCTGCGGGAGAGACCCCAAGGGAGGAGACGGGCTGGCCCTTGCCCGGAGCACC	255
Qy	915	TTCCGTCTCTAGTCCGAGTCTGAATCGGCTTTGGACCCCTGCTTGCTTCGGGGACCCC	974
Db	254	TTCCGTCTCTAGTCCGAGTCTGAATCGGCTTTGGACCCCTGCTTGCTTCGGGGACCCC	195
Qy	975	TGCAAGACGTCCACAGGCGCGCTTCCTCTCTGCTTTTATCTCTCCCGACAGCT	1034
Db	194	TGCAAGACGTCCACAGGCGCGCTTCCTCTCTGCTTTTATCTCTCCCGACAGCT	135
Qy	1035	CTGGCAGGAACCGCTCATCGTTACGCCCTTTTCGACCTCAGACCTGAGGCGGAGACC	1094
Db	134	CTGGCAGGAACCGCTCATCGTTACGCCCTTTTCGACCTCAGACCTGAGGCGGAGACC	75
Qy	1095	GCCTTGGCGCTCACTTAGACGGGACCCGGGGATGTGGCGGAGTCTCGGCTGCGCTGA	1154
Db	74	GCCTTGGCGCTCACTTAGACGGGACCCGGGGATGTGGCGGAGTCTCGGCTGCGCTGA	15
Qy	1155	CCAATCGAGTGG 1168	
Db	14	CCAATCGAGTGG 1	
RESULT 13			
ADA53808/c			
ID ADA53808 standard; cDNA; 1904 BP.			
XX	XX	ADA53808;	
XX	XX		
DT	DT	20-NOV-2003 (first entry)	
XX	XX		
DE	DE	Human coding sequence, SEQ ID 1376.	
XX	XX		
KW	KW	Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;	
KW	KW	Gene Therapy; human; secretory protein; membrane proteins; cancer;	
KW	KW	Inflammatory disease; osteoporosis; neurological disease; gene, ss.	
XX	XX		
OS	OS	Homo sapiens.	
XX	XX		
PN	PN	EF1293569-A2.	
XX	XX		
PD	PD	19-MAR-2003.	
XX	XX		
XX	XX	21-MAR-2002; 2002EP-00006586.	
PF	PF		
XX	XX		
PR	PR	14-SEP-2001; 2001JP-00328381.	
FR	FR	24-JAN-2002; 2002US-0350435P.	
XX	XX		
PA	PA	(HELI-) HELIX RES INST.	
PA	PA	(REAS-) RES ASSOC BIOTECHNOLOGY.	

XX Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,  
PI Yamamoto J, Iseno Y, Hio Y, Otsuka K, Nagai K, Irie R, Fumechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuko Y;  
XX  
XX WPI; 2003-395539/38.  
DR P-PSDB; ADA55447.  
DR  
XX  
XX New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases in



Db 237 CACTGATACAGCTGCTCTCCCATCTCTCACACAGTCCACTCCACCGGAGAGTG 296  
QY 1772 AGCCAGAGGACAGGACAGAGCTGCTGCTTTCTCTGTGTGCCAGGCGCCAGCAAGGG 1831  
Db 297 AGCCAGAGGACAGGACAGAGCTGCTGCTTTCTCTGTGTGCCAGGCGCCAGCAAGAG 356  
QY 1832 AATGTAGGAGGCTGGAGGTCAGGCGAGCTGGGATTAGGGTTGAGGGTGGGCTGTTG 1891  
Db 357 AATGTAGGAGGCTGGAGGTCAGGCGAGCTGGGATTAGGGTTGAGGGTGGGCTGTTG 416  
QY 1892 GAGGCTGAGTCTGATCTCTGCTTTAGTGGAAAGTGTCCCTTTAAACAGCAACTGGCCTGGCC 1951  
Db 417 GAGG-----CTGGATCTCTGCTTTAGTGGAAAGTGTCCCTTTAAACAGCGGCTGGCGTGGCC 470  
QY 1952 TGGCTCGGGCCCTGCTTTGGCTCCTCTGTTCACCTGGGGC 1989  
Db 471 TGGCTCGGGCCCTGCTTTGGCTCCTCTGTTCACCTGGGGC 508

RESULT 15  
AAH87257/C  
ID AAH87257 standard; DNA; 268 BP.  
XX AC AAH87257;  
XX AC  
DT 27-FEB-2002 (first entry)  
XX  
DE Human single nucleotide polymorphism containing DNA sequence #2114.  
XX  
XX Biallelic marker; polymorphism; human; disease; diagnosis; treatment;  
KW phenotypic trait; gene therapy; forensic; paternity; mapping; cancer;  
KW transgenic; single nucleotide polymorphism; SNP; ss.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT variation replace(60,A)  
FT /\*tag= a  
FT /standard\_name= "single nucleotide polymorphism"  
XX  
XX WO953095-A2.  
XX  
XX 21-OCT-1999.  
XX  
XX 30-MAR-1999; 99WO-US006893.  
XX  
XX 09-APR-1998; 98US-00057871.  
XX  
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
XX  
XX Lander ES, Wang D, Hudson T;  
XX  
XX WPI; 1999-620443/53.  
XX

PT Polymorphic human genomic sequences and related allele-specific probes  
PT and primers, useful for genetic analysis, e.g. diagnosis and monitoring  
PT of disease.  
XX  
XX  
XX Claim 1; Page 263; 330pp; English.  
XX  
XX This invention describes novel human nucleic acid segments (I) containing  
XX polymorphic sites. The polynucleotides of (I) are used for, e.g.  
XX correlating disease polymorphisms (or disease susceptibility) or other  
XX phenotypic traits (e.g. baldness, obesity, fertility, strength, response  
XX to drugs etc.); diagnosing and monitoring e.g. cancer, inflammation,  
XX heart or central nervous system diseases; detecting susceptibility to  
XX microbial infection; treating or preventing such diseases; forensic  
XX analysis; gene therapy; paternity testing; mapping genomic loci  
XX associated with phenotypic traits (and subsequent cloning of the genes  
XX responsible); and the production of transgenic organisms. Antibodies  
XX raised against (I) are useful as diagnostic and therapeutic tools and in  
XX drug screening. AAH85144 - AAH87644 represent the human DNA sequences

CC containing biallelic polymorphic sites described in the invention  
XX  
SQ Sequence 268 BP; 75 A; 55 C; 56 G; 82 T; 0 U; 0 Other;  
  
Query Match 12.8%; Score 257; DB 2; Length 268;  
Best Local Similarity 99.6%; Pred. No. 4.2e-58;  
Matches 268; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
  
QY 364 GTGAGCTATGATCCTGCCACTGCACCTGAAAGCTGGGTGACAGCAAGACCCCTGGCTCTA 423  
Db 268 GTGAGCTATGATCCTGCCACTGCACCTGAAAGCTGGGTGACAGCAAGACCCCTGGCTCTA 209  
QY 424 ATAAATGAATACATAAAGTCTCACAGCTAGTGGTAGCTAATCTGCGCAGAGTCAGGCCTC 483  
Db 208 ATAAATGAATACATAAAGTCTCACAGCTAGTGGTAGCTAATCTGCGCAGAGTCAGGCCTC 149  
QY 484 TACCTGCTGATGACAAATGGCACACTATGTCTTTTAACTGATTGCAGACCAAAATGT 543  
Db 148 TACCTGCTGATGACAAATGGCACACTATGTCTTTTAACTGATTGCAGACCAAAATGT 89  
QY 544 TTTGTGAATATTTTCCCGAGGGAATAAACCGGAAGTAGTTCCTAAATTTCTATCATCATT 603  
Db 88 TTTGTGAATATTTTCCCGAGGGAATAAACCGGAAGTAGTTCCTAAATTTCTATCATCATT 30  
QY 604 ATATTAGTTTACCTGTGGATTGGGAAAA 632  
Db 29 ATATTAGTTTACCTGTGGATTGGGAAAA 1

Search completed: November 21, 2004, 18:28:54  
Job time : 971.553 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 3207.1 Seconds  
(without alignments)  
11373.559 Million cell updates/sec

Title: US-09-820-095B-3\_COPY\_10000\_11000  
Perfect score: 1001  
Sequence: 1 acacagtcacottcgcaag.....tgggttttcacaatttggc 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues  
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_htc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	154.8	15.5	2191	3	BC047287 Homo sapi
2	154.4	15.4	353	2	AW451665 UI-H-Bi3-
3	153.4	15.3	733	4	BI753477 603026391
4	152.6	15.2	292	1	AI703080 wd81f03.x
5	152.4	15.2	525	8	B56395 CIT-HSP-200
6	152.2	15.2	414	5	BX504814 DKFZp686J
7	152.2	15.2	681	9	AG059520 Pan trogl
8	151.6	15.1	954	5	BQ231446 AGENCOURT
9	151.2	15.1	1150	4	BM807563 AGENCOURT
10	151.1	15.1	618	2	BF680805 602156044
11	150.8	15.1	322	1	AA531580 nj65a07.s
12	150.8	15.1	664	5	BQ005623 UI-H-E20-
13	150.6	15.0	803	8	BZ601931 WHAD092TR
14	150.4	15.0	535	7	CR558426 DKFZp459J
15	150.2	15.0	541	7	CN263446 170004240
16	150.2	15.0	338	2	AW023111 df49d08.Y
17	149.8	15.0	453	8	AQ285718 RPI11-87
18	149.2	14.9	794	1	AI908575 RC-BT177-
19	149.2	14.9	807	5	BUB54654 AGENCOURT
20	149.1	14.9	855	8	AQ275774 CITBI-E1-
21	149.0	14.9	868	2	BF337291 602034658
22	148.8	14.9	379	8	AQ083206 RPI11-54
23	148.8	14.9	631	8	AQ382985 RPI11-13
24	148.8	14.9	2332	3	CR601647 full-leng

C 25	148.6	14.8	669	9	AG182560	AG182560 Pan trogl
C 26	148.4	14.8	7182	3	HSMB08434	EX648286 Homo sapi
C 27	148.2	14.8	597	8	AQ035555	AQ035555 CIT-HSP-2
C 28	148	14.8	383	1	A584125	A584125 NO10G08.5
C 29	148	14.8	444	8	AQ076419	AQ076419 CIT-HSP-2
C 30	147.8	14.8	700	7	CN412525	CN412525 170005325
C 31	147.6	14.7	529	4	BG618537	BG618537 602645441
C 32	147.6	14.7	554	4	BG546209	BG546209 602573433
C 33	147.6	14.7	627	5	BU675707	BU675707 UI-CF-DU1
C 34	147.6	14.7	693	9	AG115646	AG115646 Pan trogl
C 35	147.6	14.7	851	6	CB958913	CB958913 AGENCOURT
C 36	147.6	14.7	930	5	BQ710436	BQ710436 AGENCOURT
C 37	147.4	14.7	484	8	AQ275681	AQ275681 CITBI-E1-
C 38	147.4	14.7	787	6	CB310629	CB310629 AGENCOURT
C 39	147.2	14.7	489	8	AQ190713	AQ190713 HS 3231 B
C 40	147.2	14.7	507	7	CN360588	CN360588 176005325
C 41	147.2	14.7	576	4	BM709044	BM709044 UI-E-CQ1-
C 42	147.2	14.7	585	6	CD710540	CD710540 EST-2067
C 43	147.2	14.7	720	8	BZ772985	BZ772985 mcv59H03.
C 44	147.2	14.7	762	4	BG562992	BG562992 602581160
C 45	147.2	14.7	771	4	BG624056	BG624056 602648246

ALIGNMENTS

RESULT 1  
BC047287  
LOCUS  
DEFINITION Homo sapiens, purinergic receptor P2X-like 1, orphan receptor,  
clone IMAGE:5198937, mRNA.  
ACCESSION BC047287  
VERSION BC047287.1  
KEYWORDS GI:28703829  
SOURCE HTc.  
ORGANISM Homo sapiens (human)  
REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2191)  
Strausberg, R.  
Direct Submission  
Submitted (28-FEB-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Sequencing Center (NISC),  
Gaithersburg, Maryland  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Akhter, N., Ayele, K., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,  
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., McGaspri, R.,  
Maduro, Q., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,  
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 93 Row: 1 Column: 7  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4885534  
This clone has the following problem: frame shifted.  
FEATURES  
Location/Qualifiers  
1..2191

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="NCI CGAP Sub5"
/tissue_type="Brain, adult, 6 pooled whole brains"
/clone_lib="NIH MGC_114"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"

```

## ORIGIN

```

Query Match      15.5%; Score 154.8; DB 3; Length 2191;
Best Local Similarity 71.0%; Pred. No. 1e-19;
Matches 281; Conservative 0; Mismatches 22; Indels 93; Gaps 2;

QY 107 CCAGTCCCAATGCCCTTGAGACCTGGGACCCACCTATTTTAAGCACTGGCGCTATGAAC 166
DB 618 CTAAGTCCCAATGCCCTTGAGACCTGGGACCCACCTATTTTAAGCACTGGCGCTATGAAC 677
QY 167 CACAATTCAGCCCTTACTGTCCCGTGTTCGCACTGGGACCTCGTGGCCAAAGCTGGAG 226
DB 678 CACAATTCAGCCCTTACTGTCCCGTGTTCGCACTGGGACCTCGTGGCCAAAGCTGGAG 737
QY 227 GGACCTTCGAGACCTGGCGTGTGCTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGG 286
DB 738 GGACCTTCGAGACCTGGCGTGTGCT----- 762
QY 287 GCTCTGGGAGAGGGTCCCGGGCCACCCACCGGTGGMAAGCTATGTCTATGTGACGGG 346
DB 763 -----GGG 765
QY 347 TGGCTCTAGGATCAGATTCAGTCTGGATTTGACCTGACACCGGGGACTCTGGCTG 406
DB 766 TGGCTCTAGGATCAGATTCAGTCTGGATTTGACCTGACACCGGGGACTCTGGCTG 825
QY 407 CTGGGCTCACTACTCTCTCCAGCTGACGAGAGAGCTACAACTTCAGGTGAGGC-CCCA 465
DB 826 CTGGGCTCACTACTCTCTCCAGCTGACGAGAGAGCTACAACTTCAGGTGAGGC-CCCA 885
QY 466 CTGGTCCCAAGTCCCGAGTGTGGGCCCATCGCCCT 501
DB 886 CTGGTGGGAGCAACCGGGTGTGGAGCCCGCACCT 921

```

## RESULT 2

```

AW451665      353 bp mRNA linear EST 17-FEB-2000
LOCUS         UI-H-B13-alj-f-10-0-UI.s1 NCI CGAP Sub5 Homo sapiens cDNA clone
DEFINITION    IMAGE:2736931.3, mRNA sequence.
ACCESSION     AW451665
VERSION       AW451665.1 GI:6992441
KEYWORDS      ESI.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
               1 (bases 1 to 353)
REFERENCE     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE         Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgapsb@mail.nih.gov
               Oligo-dT track not found, Not 1 site shown in beginning of sequence
               is likely internal to the message. cDNA Library Preparation: M.B.
               Soares Lab Clone distribution: NCI-CGAP clone distribution
               Information can be found through the I.M.A.G.E. Consortium/LLNL at:
               www-bio.llnl.gov/bbrp/image/image.html
               Seq primer: M13 Forward
               POLYA=No.
FEATURES      Location/Qualifiers
               1..353
               /organism="Homo sapiens"
               /mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:2736931"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Sub5"
/note="Vector: pTT3-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; NCI CGAP Sub5
is a subtracted library derived from NCI CGAP Sub4. The
NCI CGAP Sub5 library had 3 million recombinants. A
single-stranded DNA preparation of NCI CGAP Sub4 was used
as a tracer in a subtractive hybridization with a driver
comprising: the IMAGE pool (NCI CGAP Kid3 pool 1 LLAM
3334-3337, 3682-3683, 3798-3803 (IMAGE Clones);
1322376-1323911, 1456008-1456775 (1500552-1502855);
NCI CGAP Kids pool 1 LLAM 3338-3342, 3722-3725, 3776-3778
(IMAGE Clones); NCI CGAP Lys pool 1 LLAM 3573-3582,
1492104-1493255); NCI CGAP Lys pool 1 LLAM 3573-3582,
3851-3854 (IMAGE Clones);
1414920-1417991, 1520904-1522439); NCI CGAP GC4 pool 1 LLAM
3164-3167, 3716-3720, 3733-3735 (IMAGE Clones);
1257096-1258631, 1469064-1470983, 1475592-1476743);
NCI CGAP Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068
(IMAGE Clones); NCI CGAP Co10 pool 1 LLAM 2644-2653,
1217928-1220615); NCI CGAP Co10 pool 1 LLAM 2644-2653,
2871-2872 (IMAGE Clones);
1057416-1061255, 1144584-1145351). (10% of the driver
population), plus a pool of 3,840 arrayed clones from
NCI CGAP Sub1 (IMAGE Clones; 2708616-2710535) and
NCI CGAP Sub2 (IMAGE Clones; 2710536-2712435) (10% of
the driver population), plus a pool of 11,136 clones from
NCI CGAP Sub3 (IMAGE Clones; 2712456-2723591) (10% of the
driver population), plus a pool of 5,472 clones from
NCI CGAP Sub4 (IMAGE Clones; 2723592-2728969) (70% of the
driver population). Subtraction was performed as
previously described (Bonaldo, Lennon & Soares (1996):
Normalization and Subtraction: Two Approaches To
Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG TISSUE=chronic lymphocytic leukemia
TAG_LIB=NCI CGAP_CLL1
TAG_SEQ=ATTGC"

```

## ORIGIN

```

Query Match      15.4%; Score 154.4; DB 2; Length 353;
Best Local Similarity 71.1%; Pred. No. 1.5e-19;
Matches 280; Conservative 0; Mismatches 21; Indels 93; Gaps 2;

QY 109 AGTCCCAATGCCCTTGAGACCTGGGACCCACCTATTTTAAGCACTGGCGCTATGAACA 168
DB 18 AGTCCCAATGCCCTTGAGACCTGGGACCCACCTATTTTAAGCACTGGCGCTATGAACA 77
QY 169 CAATTCAGCCCTTACTGTCCCGTGTTCGCACTGGGACCTCGTGGCCAAAGCTGGAGG 228
DB 78 CAATTCAGCCCTTACTGTCCCGTGTTCGCACTGGGACCTCGTGGCCAAAGCTGGAGG 137
QY 229 ACCTTCAGGACCTGGCGTGTGCTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGC 288
DB 138 ACCTTCAGGACCTGGCGTGTGCT----- 160
QY 289 TCTGGGAGAGGGTCCCGGGCCACCCACCGGTGGAAAAGCTATGTGTATGTGCGAGGTG 348
DB 161 -----GGGTG 165
QY 349 GCTCTAGGCATCAGATTCAGTCTGGATTGTGACCTGGACACCGGGGACTCTGGCTGCT 408
DB 166 GCTCTAGGCATCAGATTCAGTCTGGATTGTGACCTGGACACCGGGGACTCTGGCTGCT 225
QY 409 GGCTCACTACTCTCTCCAGCTGCAGGAGAAGAGCTTACAACTTCAGGTGAGGC-CCCAC 467
DB 226 GGCTCACTACTCTCTCCAGCTGCAGGAGAAGAGCTTACAACTTCAGGACAGCCACTCACT 285
QY 468 GCTCCAGTGCAGTGTGCTGGGCCCATCGCCCT 501
DB 286 GGTGGGAGCAACCGGGTGTGGAGCCCGCACCT 319

```

639 CTGGCCTCACTACTCTTCAGCTTCAGGAGAGAGAGCTACACTTCAGGACAGCCAC 695

Db

RESULT 4  
LOCUS AI703080  
DEFINITION w891f03.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2338013 3' similar to SW:P2X6\_HUMAN O15547 P2X PURINOCEPTOR 6 ; mRNA sequence.  
ACCESSION AI703080.1 GI:4990980  
VERSION AI703080  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 292)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
Insert length: 800 Std Error: 0.00  
Seq primer: -40UP from Gibco.  
FEATURES  
source  
1..292  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2338013"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Lu24"  
/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-152439). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN  
Query Match 15.2%; Score 152.6; DB 1; Length 292;  
Best Local Similarity 72.8%; Pred. No. 3.5e-19;  
Matches 25; Conservative 0; Mismatches 4; Indels 92; Gaps 1;  
Qy 111 GTCCAAATGCTTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCGCTATGAACACCA 170  
Db 2 GTCCAAATGCTTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCGCTATGAACACCA 61  
Qy 171 ATTCAGCCCTACTGTCCCGTGTTCGCATTTGGGACCTCGTGGCCAGGCTGAGGAC 230  
Db 62 ATTCAGCCCTACTGTCCCGTGTTCGCATTTGGGACCTCGTGGCCAGGCTGAGGAC 121  
Qy 231 CTTTCAGGACCTCGCGTGTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGGCTC 290  
Db 122 CTTTCAGGACCTCGCGTGTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGGCTC 142  
Qy 291 TGGGAGAGGGTCCCGGGCCACCACCGGTGGAAAGCTATGTGTATGTGCGAGGTGGC 350  
Db 143 -----GGGTGGC 149

603026391F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5196695 5', mRNA sequence.  
BI753477 733 bp mRNA linear EST 25-SEP-2001  
LOCUS BI753477  
DEFINITION NIH\_MGC http://mgi.nci.nih.gov/  
ACCESSION BI753477.1 GI:15745055  
VERSION BI753477  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 733)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLAM11492 row: f column: 24  
High quality sequence stop: 719.  
FEATURES  
source  
1..733  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5196695"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_114"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 Kb, insert size range 1-3 Kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 15.3%; Score 153.4; DB 4; Length 733;  
Best Local Similarity 72.5%; Pred. No. 2.2e-19;  
Matches 259; Conservative 0; Mismatches 6; Indels 92; Gaps 1;  
Qy 107 CCAGGTCCAATGCTTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCGCTATGAAC 166  
Db 431 CTAAGTCCAATGCTTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCGCTATGAAC 490  
Qy 167 CACAATTGAGCCCTACTGTCGGTGTTCGGATTGGGACCTGTCGCCAAGGCTGGAG 226  
Db 491 CACAATTGAGCCCTACTGTCGGTGTTCGGATTGGGACCTGTCGCCAAGGCTGGAG 550  
Qy 227 GGACCTTCGAGGACCTGGGCTGTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGG 286  
Db 551 GGACCTTCGAGGACCTGGGCTGTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGG 575  
Qy 287 GCTCTGGGAGAGGGTCCCGGGCCACCACCGGTGGAAAGCTATGTCTATGTGCGAGGG 346  
Db 576 -----GGG 578  
Qy 347 TGGCTCTGTAGGCATCAGAGTTCACTGGGATTGTGACCTGGACACCGGGACTCTGGCTG 406  
Db 579 TGGCTCTGTAGGCATCAGAGTTCACTGGGATTGTGACCTGGACACCGGGACTCTGGCTG 638  
Qy 407 CTGGCCTCACTACTCTTCAGCTTCAGGAGAGAGCTACAACTTCAGGTGAGGCC 463

```

QY 351 TCTGTAGGCATCAGAGTTTCACTGGGATTGTGACCTGGACACCGGGGACTCTGGCTGCTGG 410
Db 150 TCTGTAGGCATCAGAGTTTCACTGGGATTGTGACCTGGACACCGGGGACTCTGGCTGCTGG 209
QY 411 CTTCACTACTCTCTCCAGCTGACGAGAGAGCTACAACTTCAGGTGAGGCC 463
Db 210 CTTCACTACTCTCTCCAGCTGACGAGAGAGCTACAACTTCAGGTGAGGCC 262

RESULT 5
B56395
LOCUS B56395 525 bp DNA linear GSS 20-JUN-1998
DEFINITION CIT-HSP-2006M20.TR CIT-HSP Homo sapiens genomic clone 2006M20,
genomic survey sequence.
ACCESSION B56395
VERSION B56395.1 GI:2610729
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 525)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building
JOURNAL Unpublished (1997)
COMMENT Other GSSs: CIT-HSP-2006M20.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
1..525
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7039798"
/db_xref="taxon:9606"
/clone="2006M20"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/notes="Vector: paeloBAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN
Query Match 15.2%; Score 152.4; DB 8; Length 525;
Best Local Similarity 72.8%; Pred. No. 3.6e-19;
Matches 211; Conservative 0; Mismatches 76; Indels 3; Gaps 1;

QY 712 TCCACCCACCTCGGTTCTAGTATCTCCCTCCACAGCAATGGGGTTCATTTTACT 771
Db 177 TTCAGTGTCTCTAGTGTCTAATTTCTTAATTTTGTGTTTGTGTTTCTTCT 236
QY 772 TTCCTCTTCCCTTCAGCTTTGTTTTTTTTTTTAAAGACAGAACTCATTTCTGTAC 831
Db 237 TTTTCTTTCTTTTTTTTTTTTTTTTTTTTTTTTGGACAGAGTCTCATTTCTGTAC 296
QY 832 CCAGCTGAGTGCAGTGGCCGACCTGGGTCACTGTAACTCTGTTCTCTGGTTCAA 891
Db 297 CTTGGCTGGAGTGCAGTGGCATGATCTGGGTCACTGCAACCTCTGGCTCCCGGTTCAA 356
QY 892 CCGATTCTCCCTTCAGCTTCCTGAGTAGCTGGAATTACAGGTGCTCGCCACTACTCC 951

```

```

Db 357 GGCATTCTCTGCTCAGTCTCTGAGTAGTGGATTACAGGCATGCACCACCGCCT 416
QY 952 AGCTAATTTTATATTTTGGTAGATAGATGGGTTTTTCACAATGTTGGC 1001
Db 417 GGCTAATTTTGTATTT---TTAGTAGAGATGGGGTTTTCACCAATGTTGGC 463

RESULT 6
BX504814 414 bp mRNA linear EST 04-SEP-2003
LOCUS DKP2p666J05182_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION DKP2p666J05182.5, mRNA sequence.
ACCESSION BX504814
VERSION BX504814.1 GI:32032215
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 414)
AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
TITLE EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
Wellenreuther,R., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No si sequence available.
This clone (DKFZp666J05182) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp666J05182"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/notes="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiI;
cDNA-collection"

ORIGIN
Query Match 15.2%; Score 152.2; DB 5; Length 414;
Best Local Similarity 79.2%; Pred. No. 4e-19;
Matches 194; Conservative 0; Mismatches 48; Indels 3; Gaps 1;

QY 757 TGTTCATTTTACTTCCCTTCCTCCCTTCAGCTTTGTTTTTTTTTTTAAAGACAGA 816
Db 328 TTTTTCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 269
QY 817 ATCTCATTTCTGTACCCAGCTGTGAGTGCAGTGGCCGACCTCGCTCACTGTAACTCT 876
Db 268 GTCTTGTCTGTACCCAGCTGTGAGTGCAGTGGCCGATCTCAGCTCACTGCAGCTCC 209
QY 877 GCTTCCTGGTTTCAACCGATTCTCCTCTCAGCTCTCTGAGTCTGGAATTACAGTG 936
Db 208 ACTTCCGGTTTCAACCGATTCTCCTCTCAGCTCTCTGAGTCTGGAATTACAGTG 149
QY 937 CTCGCCACTACTCCAGCTAAATTTTATTTTGTGTAGATAGATGGGTTTTCACAAAT 996
Db 148 CCGGCCACCATACCCAGCTAAATTTTGTATTT---TTAGTAGAGATGGGATTTCCCATG 92
QY 997 TTGGC 1001

```

```

Db          91 TTGGC 87

RESULT 7
AG059520/c
LOCUS      AG059520      681 bp      DNA      linear      GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-046L20.F, genomic survey sequence.
ACCESSION  AG059520
VERSION     AG059520.1  GI:16610750
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM    Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
AUTHORS     Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE       BAC end sequences of Library PTB
JOURNAL
REFERENCE   2
AUTHORS     Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE       Direct Submission
JOURNAL
COMMENT     Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpses@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
FEATURES    Location/Qualifiers
            1..681
             /organism="Pan troglodytes"
             /mol_type="genomic DNA"
             /db_xref="taxon:9598"
             /clone="PTB-046L20.F"
             /sex="male"
             /cell_type="lymphoblast"
             /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      15.2%; Score 152.2; DB 9; Length 681;
Best Local Similarity 82.2%; Pred. No. 3.8e-19;
Matches 175; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY      788 CAGCTTTGTTTTTTTTTTTAAAGACAGAACTCTCATCTGTGACCCAGGCTGGAGTGCAG 847
Db      507 CAATTTTTTTTTTTTTTTTGTAGAGGGAGTCTCACTCTGTCCACCGAGCTGGAGTGCAA 448
QY      848 TGGCCGACCTGGCTCACTGAACCTCTGCTTCTGGGTTCAACCGATTCTCTTCCTC 907
Db      447 TGGCAAACTTGGCTCACTGCACAACTCTGCTCCAGGTTCAAGGATCTCTGCTC 388
QY      908 AGCTCTCAGTAGTGGAAATACAGGTGCTGCCACTACTCCAGCTAAATTTTATTT 967
Db      387 AGCTCCCGAGTAGCTAGGATTACAGGTGCTGCCACCATGCCAGCTAAATTTTGT 328
QY      968 TTGGTAGATAGATGGGTTTTCACAACTTGG 1000
Db      327 ATTTTAGTAGATGGGTTCTCACCATTGG 295

RESULT 8
BQ231446
LOCUS      BQ231446      954 bp      mRNA      linear      EST 02-MAY-2002

DEFINITION AGNCOURT 7589565 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6067564
            5', mRNA sequence.
ACCESSION  BQ231446
VERSION     BQ231446.1  GI:20412846
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/,
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            plate: LLAM13347 row: e column: 05
            High quality sequence stop: 576.
            Location/Qualifiers
            1..954
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6067564"
             /tissue_type="embryonal carcinoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_92"
             /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
            Average insert size 2.5 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC Library."
ORIGIN
Query Match      15.1%; Score 151.6; DB 5; Length 954;
Best Local Similarity 81.8%; Pred. No. 4.7e-19;
Matches 175; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY      788 CAGCTTTGTTTTTTTTTTTAAAGACAGAACTCTCATCTGTGACCCAGGCTGGAGTGCAG 847
Db      178 CAGACTTTTTTTTTTTTTTTTGTAGACAGAGTTTCACTCTTGCACCCAGGCTGGGTACAA 237
QY      848 TGGCCGACCTGGCTCACTGTAACTCTGCTTCTGGGTTCAACCGATTCTCTTCCTC 907
Db      238 TGGTGTGATCTTGGCCCACTGTAACCTCTGCTTCTGGGTTCAAGCATCTCTGCTC 297
QY      908 AGCTCTCAGTAGTGGAAATACAGGTGCTGCCACTACTCCAGCTAAATTTTATTT 967
Db      298 AGCTCTCAGTAGTGGGATTACAGGTGCTGCCACACACCTGGCTAAATTTTGTATT 357
QY      968 TTGGTAGATAGATGGGTTTTCACAACTTGGC 1001
Db      358 TTTTAGTAGATGCGGGGTTTCCCATGTTGGC 391

RESULT 9
BQ231446
LOCUS      BQ231446      1150 bp      mRNA      linear      EST 05-MAR-2002
DEFINITION AGNCOURT 6574865 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5732451
            5', mRNA sequence.
ACCESSION  BQ231446
VERSION     BQ231446.1  GI:19124386
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```



Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo F. Chuquai,  
 M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: David B. Krizman, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html  
 Insert length: 431 Std Error: 0.00  
 Seq primer: -40m13 fwd. Et from Amersham.

#### FEATURES

Location/Qualifiers  
 1. 322  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:997332"  
 /sex="male"  
 /tissue\_type="invasive prostate tumor"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Pr10"  
 /notes="Organ: prostate; Vector: PAMP10; mRNA made from  
 invasive prostate tumor, cDNA made by oligo-dT priming.  
 Non-directionally cloned. Size-selected on agarose gel,  
 average insert size 600 bp. Library made by D. Krizman,  
 NIH."

#### ORIGIN

Query Match 15.1%; Score 150.8; DB 1; Length 322;  
 Best Local Similarity 80.7%; Pred. No. 7.7e-19;  
 Matches 176; Conservative 0; Mismatches 42; Indels 0; Gaps 0;  
 QY 784 CCTTCAGCTTGTGTTTTTTTTTTAAAGACAGAAATCTCATTTCTGTCACCCAGGCTGGAGT 843  
 Db |||  
 QY 67 CCTCTTGGCTCAGCCTCTCTTTTTCAGACAGAGTCTGCTGTGTCACCCAGGCTGGAGT 126  
 Db |||  
 QY 844 GCAGTGGCCGACCTCGGCTCACTGTAACTCTGCTTCTGCTGGGTTCACCGATTCCTCTT 903  
 Db |||  
 QY 127 GCAGTGGCAGATCTCGGCTCACTGCAACCTCTGCTCTGCTGGTTCACGAATTCCTCG 186  
 Db |||  
 QY 904 CCTCAGCCTCCTGAGTAGCTGGAATTACAGGTGCTCGCCACTCTCCAGCTAAATTTTA 963  
 Db |||  
 QY 187 CCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCTGCCACACGCCAGCTAAATTTT 246  
 Db |||  
 QY 964 TATTTTGGTAGATAGATGGGTTTTCACAATGTGGC 1001  
 Db |||  
 QY 247 TAGTATTTTAGTAGAGAAAGGGTTTTCACCGTGTAGC 284  
 Db |||

#### RESULT 12

BQ005623  
 LOCUS  
 DEFINITION  
 UI-H-ED0-ayr-a-11-0-UI.s1 NCI\_CGAP\_ED0 Homo sapiens cDNA clone  
 IMAGE:5842450 3', mRNA sequence.  
 ACCESSION  
 BQ005623  
 EST  
 BQ005623.1 GI:19730523  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 664)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Dr. Jose Mercuende  
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 The following repetitive elements were found in this cDNA  
 sequence: 1-242, >SVA#Other 94-380, >ALU (matched complement)  
 394-606, >ALU (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

#### FEATURES

Location/Qualifiers  
 1. 664  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5842450"  
 /tissue\_type="Chondrosarcoma"  
 /dev\_stage="Adult"  
 /lab\_hosts="DH10B (Life Technologies)"  
 /clone\_lib="NCI\_CGAP\_ED0"  
 /notes="Organ: Left Pubic Bone; Vector: pT73-Pac  
 (Pharmacia) with a modified polylinker; Site 1: EcoR I;  
 Site 2: Not I; NCI\_CGAP\_ED0 is a cDNA library containing  
 the following tissue(s): Chondrosarcoma cell line CSS. The  
 library was constructed according to Bonaldo, Lennon and  
 Soares Genome Research, 6:791-806, 1996. First strand  
 cDNA synthesis was primed with an oligo-dT primer  
 containing a Not I site. Double stranded cDNA was ligated  
 to an EcoR I adaptor, digested with Not I, and cloned  
 directionally into pT73-Pac vector. The oligonucleotide  
 used to prime the synthesis of first-strand cDNA contains  
 a library tag sequence that is located between the Not I  
 site and the (dT)18 tail. The sequence tag for this  
 library is GCTCAGGCT.  
 TAG\_L1SUB=chondrosarcoma  
 TAG\_L1B=UI-H-ED0  
 TAG\_SEQ=CGTCAGGCT"

#### ORIGIN

Query Match 15.1%; Score 150.8; DB 5; Length 664;  
 Best Local Similarity 80.7%; Pred. No. 7e-19;  
 Matches 176; Conservative 0; Mismatches 42; Indels 0; Gaps 0;  
 QY 784 CCTTCAGCTTGTGTTTTTTTTTTAAAGACAGAAATCTCATTTCTGTCACCCAGGCTGGAGT 843  
 Db |||  
 QY 76 CCTCTTGGCTCAGCCTCTCTTTTTCAGACAGAGTCTGCTGTCTCACCAGGCTGGAGT 135  
 Db |||  
 QY 844 GCAGTGGCCGACCTCGGCTCACTGTAACTCTGCTTCTCTGGGTTCAACGATTCCTCTT 903  
 Db |||  
 QY 136 GCAGTGGCATGATCTCGGCTCACTGCAACCTCTGCTCTCTGGGTTCAAGCAATTCCTCTG 195  
 Db |||  
 QY 904 CCTCAGCCTCCTGAGTAGCTGGAATTACAGGTGCTCGCCACTCTCCAGCTAAATTTTA 963  
 Db |||  
 QY 196 CCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCTGCCACCGCCAGCTAAATTTT 255  
 Db |||  
 QY 964 TATTTTGGTAGATAGATGGGTTTTCACAATGTGGC 1001  
 Db |||  
 QY 256 TAGTATTTTAGTAGAGAAAGGGTTTTCACCGTGTAGC 293  
 Db |||

#### RESULT 13

BZ601931/c  
 LOCUS  
 DEFINITION  
 WHADO92TR Human MCF7 breast cancer cell line library (MCF7.1) Homo  
 sapiens genomic clone MCF7.1-22P15, genomic survey sequence.  
 ACCESSION  
 BZ601931  
 VERSION  
 BZ601931.1 GI:31510393  
 KEYWORDS  
 GSS.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 803)  
 Volik, S., Zhao, S., Chin, K., Brenner, J. H., Herndon, D. R., Tao, Q.,  
 Kowber, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,  
 Gray, J. W., and Collins, C.  
 End-sequence profiling: Sequence-based analysis of aberrant genomes





```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived_from H1, H7 and H9 cells"
/clone_lib="CRN_EB"
/notes="Oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."
```

## ORIGIN

Query Match	15.0%;	Score 150.2;	DB 7;	Length 541;
Best Local Similarity	83.7%;	Pred. No. 9.4e-19;		
Matches 170;	Conservative 0;	Mismatches 33;	Indels 0;	Gaps 0;
Qy	799	TTTTTTTTTAAAGACAGAACTCTATTCTGTCCACCGAGCTGGAGTGCAGTGGCCCGACCT	858	
Db	1	TTTTTTTTTTCAGACAGGCTTACTCTGTACCCAGGCTGGAGTGCAGTGGCGTGATCT	60	
Qy	859	CGGCTCACTGTAACCTCTGCTTCTGGTTCAAACCGATTCTCTCTCAGCCCTCCAG	918	
Db	61	CGSCTCACTGCAACCTCTGCCCTCTGGGTTCAAGTGATTCTCTGSCCTCAGCTCTCCCTGAG	120	
Qy	919	TAGCTGGAATTTACAGGTGCTCGCCACTACTCCCAAGTAAATTTTTATTTTGGTAGATAG	978	
Db	121	TAGCTGGGATTCAGGTTGCACACCACTGCCCGGCTAAATTTTTTTGTTATTTTAGTAG	180	
Qy	979	AGATGGGTTTTCAAAATGTTGGC	1001	
Db	181	AGACGGGATTTTCAACATGTTGGC	203	

Search completed: November 21, 2004, 17:35:13  
Job time : 3212.1 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 17:43:16 ; Search time 494.874 Seconds  
(without alignments)  
10930.100 Million cell updates/sec

Title: US-09-820-095B-3\_COPY\_10000\_11000

Perfect score: 1001

Sequence: 1 acacagtcaccttcagcaag.....tgggtttcaccaatgtggc 1001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	16449	10	US-09-820-095-3
2	393	39.3	576	9	Sequence 3, Appli
3	364	36.4	440	9	Sequence 9249, Ap
4	158.2	15.8	174448	13	US-09-864-761-9249
5	155.8	15.6	301692	16	US-10-087-192-148
6	155.8	15.6	310268	17	US-10-428-487-11
7	155.6	15.5	63045	17	US-10-367-094-195
8	155.6	15.5	63824	15	US-10-714-796-76
9	155.6	15.5	63824	15	US-10-282-174-348
10	155.6	15.5	202100	15	US-10-282-174-348
11	154.8	15.5	2693	10	US-09-820-095-1
12	153.8	15.4	14176	9	US-09-764-864-1644

C	13	153.4	15.3	6248	15	US-10-242-355-705	Sequence 705, App
	14	153	15.3	128779	15	US-10-081-327-38	Sequence 38, Appl
	15	152.6	15.2	588	13	US-10-027-632-77218	Sequence 77218, A
	16	152.6	15.2	588	13	US-10-027-632-300376	Sequence 300376, A
	17	152.6	15.2	588	13	US-10-027-632-77218	Sequence 77218, A
	18	152.6	15.2	588	13	US-10-027-632-300376	Sequence 300376, A
	19	152.2	15.2	32221	9	US-09-764-878-377	Sequence 377, App
C	20	152.2	15.2	32221	14	US-10-079-854-377	Sequence 377, App
	21	152	15.2	527	13	US-10-027-632-130193	Sequence 130193, A
	22	152	15.2	527	13	US-10-027-632-130193	Sequence 130193, A
	23	152	15.2	1364	13	US-10-027-632-253464	Sequence 253464, A
	24	152	15.2	1364	13	US-10-027-632-253464	Sequence 253464, A
	25	152	15.2	1364	13	US-10-027-632-253464	Sequence 253464, A
	26	152	15.2	1364	13	US-10-027-632-253465	Sequence 253465, A
	27	151.8	15.2	631	13	US-10-027-632-57903	Sequence 57903, A
	28	151.8	15.2	631	13	US-10-027-632-57903	Sequence 57903, A
	29	151.8	15.2	631	13	US-10-027-632-57903	Sequence 57903, A
	30	151.8	15.2	631	13	US-10-027-632-79939	Sequence 79939, A
	31	151.8	15.2	649	13	US-10-027-632-6071	Sequence 6071, Ap
	32	151.8	15.2	649	13	US-10-027-632-6071	Sequence 6071, Ap
	33	151.8	15.2	32767	15	US-10-004-113-4	Sequence 4, Appli
	34	151.8	15.2	32767	16	US-10-394-948-4	Sequence 4, Appli
	35	151.2	15.1	43009	13	US-10-087-192-1582	Sequence 1582, Ap
	36	151	15.1	32189	9	US-09-764-878-379	Sequence 379, App
	37	151	15.1	32189	14	US-10-079-854-379	Sequence 379, App
	38	150.8	15.1	529	13	US-10-027-632-267212	Sequence 267212, A
	39	150.8	15.1	529	13	US-10-027-632-267212	Sequence 267212, A
	40	150.8	15.1	98546	13	US-10-087-192-1414	Sequence 1414, Ap
	41	150.8	15.1	107543	17	US-10-322-281-706	Sequence 706, App
	42	150.8	15.1	147309	9	US-09-742-312-3	Sequence 3, Appli
	43	150.8	15.1	147309	15	US-10-436-185-3	Sequence 3, Appli
	44	150.8	15.1	201143	16	US-10-240-425-1099	Sequence 1099, Ap
	45	150.6	15.0	113000	15	US-10-376-566-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1  
US-09-820-095-3  
; Sequence 3, Application US/09820095  
; Publication No. US20030233668A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: C1001202  
; CURRENT APPLICATION NUMBER: US/09/820,095  
; CURRENT FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 16449  
; TYPE: DNA  
; ORGANISM: Human  
US-09-820-095-3

Query Match	100.0%;	Score	1001;	DB	10;	Length	16449;
Best Local Similarity	100.0%;	Pred. No.	1.3e-282;				
Matches	1001;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0;
Qy	1	ACACAGTCACCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCCC	60				
Db	10000	ACACAGTCACCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCCC	10059				
Qy	61	CAGACCCCTCTGTGCTCCCTACCTCATCTGACCTTTCCCACTCTCCAGGTCCAATGCC	120				
Db	10060	CAAGACCCCTCTGTGCTCCCTACCTCATCTGACCTTTCCCACTCTCCAGGTCCAATGCC	10119				
Qy	121	TTGAGACCTGGGACCCACCTATTTTATGACACTGCGGTATGACACCAATTCAGCCCC	180				
Db	10120	TTGAGACCTGGGACCCACCTATTTTATGACACTGCGGTATGACACCAATTCAGCCCC	10179				

181 TACTGTCCCGTGTTCGCGATTTGGGCACTCGTGTGCCCAAGGCTGGAGGACCTTCGAGGAC 240  
10180 TACTGTCCCGTGTTCGCGATTTGGGCACTCGTGTGCCCAAGGCTGGAGGACCTTCGAGGAC 10239  
241 CTGGCGTTGCTGTGGGTGCTCCCAAGTTGGGGGACAGGTTCTAGAGGCTGTGGAGAGGG 300  
10240 CTGGCGTTGCTGTGGGTGCTCCCAAGTTGGGGGACAGGTTCTAGAGGCTGTGGAGAGGG 10299  
301 TCCCGGGCCCAACCCACCGTGGAAAGCTATGCTATGTGCAGGCTGGCTCTAGGCA 360  
10300 TCCCGGGCCCAACCCACCGTGGAAAGCTATGCTATGTGCAGGCTGGCTCTAGGCA 10359  
361 TCAGAGTTTCACTGGGATTTGACCTTGGACACCGGGGACTTGGGTGTGGCTCACTACT 420  
10360 TCAGAGTTTCACTGGGATTTGACCTTGGACACCGGGGACTTGGGTGTGGCTCACTACT 10419  
421 CTTCCAGCTGAGAGAGAGCTTCAACTTCAGTGGAGCCCACTGCTCCAGTGGCC 480  
10420 CTTCCAGCTGAGAGAGAGCTTCAACTTCAGTGGAGCCCACTGCTCCAGTGGCC 10479  
481 AGCTGCTGGGCCATCGCCCTCTCACTGTGGCGGACAGACACACCCAGGCCCGAG 540  
10480 AGCTGCTGGGCCATCGCCCTCTCACTGTGGCGGACAGACACACCCAGGCCCGAG 10539  
541 GCCTCTAGATATTCATACGTGTGCAAGGGGTTCCAGGAGCAGGAGAGCTGTCTC 600  
10540 GCCTCTAGATATTCATACGTGTGCAAGGGGTTCCAGGAGCAGGAGAGCTGTCTC 10599  
601 RACCCCAATCTCCAGACAGGCTCGCTGTGCTGCCCAAGCTCTGAGCCCTCCACCC 660  
10600 RACCCCAATCTCCAGACAGGCTCGCTGTGCTGCCCAAGCTCTGAGCCCTCCACCC 10659  
661 CATCTGTCCAGGCCCTGCCAGCTCAGGCTCTCACTGCAGCCCTTCTCCACCCCA 720  
10660 CATCTGTCCAGGCCCTGCCAGCTCAGGCTCTCACTGCAGCCCTTCTCCACCCCA 10719  
721 CTTCTGTCTATGATCTCCCTTCCACAGCAATGGGGTGTTCATTTTACTTCCCTTC 780  
10720 CTTCTGTCTATGATCTCCCTTCCACAGCAATGGGGTGTTCATTTTACTTCCCTTC 10779  
781 TCCCTTCAGCTTCTGTTTTTTTTTTTTTTTTTTTAAAGACAACTCATCTGTCACCCAGGCTGG 840  
10780 TCCCTTCAGCTTCTGTTTTTTTTTTTTTTTTTTTAAAGACAACTCATCTGTCACCCAGGCTGG 10839  
841 AGTGAGTGGCCGACCTCGGCTCACTGTAACTCTGCTTCTGGGTTCAACCGATCTC 900  
10840 AGTGAGTGGCCGACCTCGGCTCACTGTAACTCTGCTTCTGGGTTCAACCGATCTC 10899  
901 CTTCTCAGCTTCTGAGTGTGAATACAGTGTCTCGCCTACTCCAGCTACTTT 960  
10900 CTTCTCAGCTTCTGAGTGTGAATACAGTGTCTCGCCTACTCCAGCTACTTT 10959  
961 TTATATTTTGGTAGATAGATGGGTTTTTCAAAATGTTGGC 1001  
10960 TTATATTTTGGTAGATAGATGGGTTTTTCAAAATGTTGGC 11000

RESULT 2  
US-09-864-761-9249/c  
; Sequence 9249, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aomica-x-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 9249  
LENGTH: 576  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC002472.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.1  
US-09-864-761-9249

Query Match 39.3%; Score 393; DB 9; Length 576;  
Best Local Similarity 100.0%; Pred. No. 8.6e-105; Indels 0; Gaps 0;  
Matches 393; Conservative 0; Mismatches 0

QY 1 ACACAGTCACTTTCAGCAAGTTTCAACTTCTCTAAGTAAGCAGAGTGGTCTCATCTGCC 60  
DB 393 ACACAGTCACTTTCAGCAAGTTTCAACTTCTCTAAGTAAGCAGAGTGGTCTCATCTGCC 334  
QY 61 CAAGAGCCCTTGTCTCCCTTACTCTCATCTGACCTTCCACCTCCCTCCAGGTTCCAATGCC 120  
DB 333 CAAGAGCCCTTGTCTCCCTTACTCTCATCTGACCTTCCACCTCCCTCCAGGTTCCAATGCC 274  
QY 121 TTGGAGACCTGGGACCCCACTTATTTTAAAGCACTGCCGCTATGAACCAACAATTCAGCCCC 180  
DB 273 TTGGAGACCTGGGACCCCACTTATTTTAAAGCACTGCCGCTATGAACCAACAATTCAGCCCC 214  
QY 181 TACTGTCCCGTGTTCGCGATTTGGGCACTCGTGTGCCCAAGGCTGGAGGACCTTCGAGGAC 240  
DB 213 TACTGTCCCGTGTTCGCGATTTGGGCACTCGTGTGCCCAAGGCTGGAGGACCTTCGAGGAC 154  
QY 241 CTGGCGTTGCTGTGGGTGCTCCCAAGTTGGGGGACAGGTTCTTAGAGGCTCTCGGAGAGGG 300

Db 153 CTGGCGTCTCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTTAGAGGGGCTCTGGGAGAGGG 94  
QY 301 TCCCGGGCCACCCACCGGTGAAAGCTATGTCTATGTGTCAGGGTGGCTCTGTAGGCA 360  
Db 93 TCCCGGGCCACCCACCGGTGAAAGCTATGTCTATGTGTCAGGGTGGCTCTGTAGGCA 34  
QY 361 TCAGAGTTCACTGGGATTTGACCTGGACACCG 393  
Db 33 TCAGAGTTCACTGGGATTTGACCTGGACACCG 1  
RESULT 3  
US-09-864-761-2179/c  
; Sequence 2179, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 2179  
; LENGTH: 440  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC002472.3  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.9  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.2

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.6  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.5  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.6  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.7  
US-09-864-761-2179  
Query Match 36.4%; Score 364; DB 9; Length 440;  
Best Local Similarity 100.0%; Pred. No. 2.5e-96;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACACAGTCACCTTCAGCAAGTTCACACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCCC 60  
Db 364 ACACAGTCACCTTCAGCAAGTTCACACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCCC 305  
QY 61 CAAGACCCCTCTTGTCTCCCTACCTCATCTGACCTTTTCCCACTCTCTCCAGGTCCAATGCC 120  
Db 304 CAAGACCCCTCTTGTCTCCCTACCTCATCTGACCTTTTCCCACTCTCTCCAGGTCCAATGCC 245  
QY 121 TTGGAGACCTGGGACCCACCTATTATTAAGCATCTGCGGCTATGAACACCAATTCAGCCCC 180  
Db 244 TTGGAGACCTGGGACCCACCTATTATTAAGCATCTGCGGCTATGAACACCAATTCAGCCCC 185  
QY 181 TACTGTCCCGTGTTCCTCCATTGGGGACCTCTGTGCGCAAGGCTGGAGGACCTTCGAGGAC 240  
Db 184 TACTGTCCCGTGTTCCTCCATTGGGGACCTCTGTGCGCAAGGCTGGAGGACCTTCGAGGAC 125  
QY 241 CTGCGGTTGCTGTGGTGTCCCAAGTTGGGGGCGAGGGTCTTAGAGGCTCTGGGAGAGGG 300  
Db 124 CTGCGGTTGCTGTGGTGTCCCAAGTTGGGGGCGAGGGTCTTAGAGGCTCTGGGAGAGGG 65  
QY 301 TCCCGGGCCACCCACCGGTGGAAAGCTATGTCTATGTGACAGGGTGGCTCTGTAGGCA 360  
Db 64 TCCCGGGCCACCCACCGGTGGAAAGCTATGTCTATGTGACAGGGTGGCTCTGTAGGCA 5  
QY 361 TCAG 364  
Db 4 TCAG 1  
RESULT 4  
US-10-087-192-148/c  
; Sequence 148, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 148  
; LENGTH: 174448  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(174448)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-148  
Query Match 15.8%; Score 158.2; DB 13; Length 174448;  
Best Local Similarity 68.8%; Pred. No. 3.3e-35;  
Matches 232; Conservative 0; Mismatches 103; Indels 2; Gaps 1;



Qy 940 GCACACTACCCAGCTAATTTTATATTTTGGTAGATAGAGATGGGTTTTCACAATGTTG 999  
 44153 GCCACACACCCAGCTAATTTTGTATTTT--TTAGTAGAGATGGGTTTGGCCATGTTG 44097  
 Qy 1000 GC 1001  
 Db 44096 GC 44095

```

RESULT 8
US-10-282-174-347/c
// Sequence 347, Application US/10282174
// Publication No. US20030224380A1
// GENERAL INFORMATION:
// APPLICANT: Becker, Kenneth David
// APPLICANT: Velicelebi, Gonul
// APPLICANT: Elliot, Kathryn J.
// APPLICANT: Wang, Xin
// APPLICANT: Tanzi, Rudolph E.
// APPLICANT: Bertram, Lars
// APPLICANT: Saunders, Aleister J.
// APPLICANT: Mullin, Kristina M.
// APPLICANT: Sampson, Andrew Johnson
// APPLICANT: Blacker, Deborah Lynne
// TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
// TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
// TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
// FILE REFERENCE: 37481-3308
// CURRENT APPLICATION NUMBER: US/10/282,174
// CURRENT FILING DATE: 2002-10-25
// PRIOR APPLICATION NUMBER: US 60/339,525
// PRIOR FILING DATE: 2001-10-25
// PRIOR APPLICATION NUMBER: US 60/338,010
// PRIOR FILING DATE: 2001-11-08
// PRIOR APPLICATION NUMBER: US 60/336,929
// PRIOR FILING DATE: 2001-11-08
// PRIOR APPLICATION NUMBER: US 60/338,363
// PRIOR FILING DATE: 2001-11-09
// PRIOR APPLICATION NUMBER: US 60/337,052
// PRIOR FILING DATE: 2001-12-04
// PRIOR APPLICATION NUMBER: US 60/368,919
// PRIOR FILING DATE: 2002-03-28
// NUMBER OF SEQ ID NOS: 564
// SOFTWARE: FastSeq for Windows Version 4.0
// SEQ ID NO 347
// LENGTH: 63824
// TYPE: DNA
// ORGANISM: Homo sapiens
// FEATURE:
// NAME/KEY: misc_feature
// LOCATION: (1)...(14)
// OTHER INFORMATION: N is unknown
// FEATURE:
// NAME/KEY: misc_feature
// LOCATION: (22850)...(22970)
// OTHER INFORMATION: N is unknown
// FEATURE:
// NAME/KEY: misc_feature
// LOCATION: (39442)...(39541)
// OTHER INFORMATION: N is unknown
// FEATURE:
// NAME/KEY: misc_feature
// LOCATION: (53423)...(53522)
// OTHER INFORMATION: N is unknown
// FEATURE:
// NAME/KEY: misc_feature
// LOCATION: 57620..57642
// OTHER INFORMATION: N is unknown
// FEATURE:
// NAME/KEY: misc_feature
// LOCATION: (57652)...(57751)
// OTHER INFORMATION: N is unknown
// US-10-282-174-347

```

US-10-282-174-347

Query Match	15.5%;	Score 155.6;	DB 15;	Length 63824;
Best Local Similarity	80.6%;	Pred. No. 1.4e-34;		
Matches 195;	Conservative 0;	Mismatches 44;	Indels 3;	Gaps 1;
Qy	760	TTCAATTTTACTTTCCTTCCTCCCTTCAGCTTGTGTTTTTTTTTTTAAAGACAGAAATC	819	
Db	45333	TTAATAATTACATTTCCATTCCTCAATCTTTTTTTTTTTTTTTTGTAGACGGAGTC	45274	
Qy	820	TCATTTCTGCACCGAGGCTGGAGTGCAGTGTGCGACCTCGGCTCACTGTAACTCTGCT	879	
Db	45273	TCATTTCTGCACCCAGGTTGGAGTACAGTGGCAACATCTGGCTCACTGCAACCTCTGCC	45214	
Qy	880	TCTGGGTTTCAACCGAATTCCTCTCCTCAGCGCTCTCTAGTAGCTGGAAATTCAGGTTGCTC	939	
Db	45213	TCTGGGTTTCAAGTGATTTCTCTCTGCTCAGCGCTTCGGAGTAGCTGGGATTATAGGTTGA	45154	
Qy	940	GGCACTACTCCAGCTAATTTTATTTTGGTAGATAGAGATGGGTTTTCACAAATGTTG	999	
Db	45153	GCACACACACCCAGCTAATTTTGTGTATTTT---TTAGTAGAGATGGGTTTCGCCATGTTG	45097	
Qy	1000	GC 1001		
Db	45096	GC 45095		

## RESULT 9

```

RESULTS
US-10-282-174-348/c
; Sequence 348, Application US/10282174
; Publication NO. US20030224380A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Kenneth David
; APPLICANT: Velicelebi, Gonul
; APPLICANT: Elliot, Kathryn J.
; APPLICANT: Wang, Xin
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bertram, Lars
; APPLICANT: Saunders, Aleister J.
; APPLICANT: Mullin, Kristina M.
; APPLICANT: Sampson, Andrew Johnson
; APPLICANT: Blacker, Deborah Lynne
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
; TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
; FILE REFERENCE: 37481-3308
; CURRENT APPLICATION NUMBER: US/10/282,174
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/339,525
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/338,010
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/336,929
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/338,363
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/337,052
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/368,919
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 564
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 63824
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 300
; OTHER INFORMATION: Insertion: CA following nt 299
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1152,15104,20815,35719,36738,36739,42125,45083,45887,56'
; LOCATION: 56887,58524,62661,63802

```





QY 407 CTGGCTCACTACTCTCTCCAGCTGCAGGAGAGAGCTACAACTTCAGGTAGGC-CCCA 465  
Db 775 CTGGCTCACTACTCTCTCCAGCTGCAGGAGAGAGCTACAACTTCAGGACAGCCACTCA 834  
QY 466 CTGCTCCAGTCCAGCTCTCGGCGCCATCGCCCT 501  
Db 835 CTGGTGGAGACACCGGTGTGAGGCCCGCACCT 870

RESULT 12  
US-09-764-864-1644/c  
; Sequence 1644, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1644  
; LENGTH: 14176  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-1644

Query Match 15.4%; Score 153.8; DB 9; Length 14176;  
Best Local Similarity 76.7%; Pred. No. 3.1e-34;  
Matches 188; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 757 TGTTCATTTTACTTTCCCTCTCCCTTCAGCTTCTGTTTTTTTTTAAAGACAGA 816  
Db 10305 TGGGACATCTTAATAGGCAATGATCTCCCTGGTTTTTTTTTTTGGACAGA 10246  
QY 817 ATCTCATTCGTCAACAGCTGGAGTGCAGTGGCCGACCTGCGGTCTACTGTAACCTCT 876  
Db 10245 GTCTGCTCTGTCAACAGCTGGAGTGCAGTGGTGTGATCTCAACTCACTGCAACCTCT 10186  
QY 877 GCTTCCTGGTTCAACCGATTCTCTCTCAGCTCTCCTGAGTGGTGAATACAGGTG 936  
Db 10185 GCTCTCTGGTTCAACCGATTCTCTCAGCTCTCCTGAGTGGTGAATACAGGTG 10126  
QY 937 CTGCGCACTACTCCCACTAAATTTTATATTTTGTGTAGATAGATGGTTTTTCAATG 996  
Db 10125 TGTGCCACGAGCTGGCTAAATTTTGTGTAGATAGATGGTTTTTCAATG 10066  
QY 997 TTGGC 1001  
Db 10065 TTGGC 10061

RESULT 13  
US-10-242-355-705/c  
; Sequence 705, Application US/10242355  
; Publication No. US2003023583A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC003C1  
; CURRENT APPLICATION NUMBER: US/10/242,355  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: 09/764,897  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1267  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 705  
; LENGTH: 6248  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-242-355-705

Query Match 15.3%; Score 153.4; DB 15; Length 6248;  
Best Local Similarity 78.3%; Pred. No. 3.1e-34;  
Matches 184; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 767 TTACTTTCCCTCTCTCCCTTCAGCTTCTGTTTTTTTTTAAAGACAGAACTCATCT 826  
Db 3739 TTACAGGCTCCACCACTGCTGTTTTTTTTTTTTTTTGTAGAGGAATCTCGCTCT 3680  
QY 827 GTCACCCAGGCTGGAGTGCAGTGGCCGACCTCGCTCACTGTAACCTCTCTCTCTGG 886  
Db 3679 GTGCGCCAGGCTGAGGAGCAGTGGCAATCTCGCTCACTGCAAGCTCCGCCCTCTGG 3620  
QY 887 TTCAACCAATCTCTCTCTCAGCTCTCTGAGTAGTGAATACAGGTGCTCGCCACTA 946  
Db 3619 TTACAGCAATCTCTCTCTCAGCTCTCTGAGTAGTGGAGTACAGGTGCGCCACCA 3560  
QY 947 CTCCAGCTAAATTTTATATTTTGTGTAGATAGATGGTTTTTACAAATGTTGG 1001  
Db 3559 CGCTAACTAAATTTTGTGATTTTGTGTAGACGGGTTTCCACGTGTAGC 3505

RESULT 14  
US-10-081-327-38  
; Sequence 38, Application US/10081327  
; Publication No. US20030129602A1  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Tim  
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION  
; FILE REFERENCE: UMO1523  
; CURRENT APPLICATION NUMBER: US/10/081,327  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: 06/120,592  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: 06/118,760  
; PRIOR FILING DATE: 1999-02-05  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 38  
; LENGTH: 128779  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-081-327-38

Query Match 15.3%; Score 153; DB 15; Length 128779;  
Best Local Similarity 77.2%; Pred. No. 1e-33;  
Matches 186; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 759 TTTCAATTTTACTTTCCCTCTCTCCCTTCAGCTTCTGTTTTTTTTTAAAGACAGAAT 818  
Db 94246 TATAAATATATCTTAAATTTTATTTTCTACTTTTTTTTTTTTGGAGACAAGAT 94305  
QY 819 CTCATTTCTCACCAGCTGGAGTGCAGTGGCCGACCTCGGTCTCACTGTAACCTCTGC 878  
Db 94306 CTTACTCTGTCAACCCAGGCTGAGTGAATAGACGGATCTCGGTCTCACTGCAACCTCTGC 94365



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 83.5046 Seconds  
(without alignments)  
8520.488 Million cell updates/sec

Title: US-09-820-095B-3\_COPY\_10000\_11000

Perfect score: 1001  
Sequence: 1 acacgctaccttcgcaag.....tgggttttcaaatgttggc 1001

Scoring table: IDENTITY NUC

Gapex 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/1/ina/5A-COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5B-COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A-COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B-COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/PCUTUS-COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	154.8	15.5	1293	3	US-09-381-681-1
2	154.8	15.5	1360	3	US-09-191-136-30
3	154.8	15.5	1697	3	US-09-381-681-2
4	153	15.3	128779	4	US-09-497-855A-38
5	150	15.0	13204	4	US-09-054-272-49
6	149.2	14.9	70000	4	US-09-851-896-3
7	148.6	14.8	36159	4	US-09-749-588-3
8	148.6	14.8	36159	4	US-10-135-687-3
9	148.4	14.8	48763	4	US-09-916-204-3
10	148.4	14.8	48763	4	US-10-282-048-3
11	147	14.7	8453	3	US-09-167-681-45
12	147	14.7	55298	4	US-09-491-356C-1
13	147	14.7	63000	4	US-09-780-172-18
14	146.2	14.6	72604	3	US-09-268-992-7
15	146.2	14.6	72604	3	US-09-657-474-7
16	145.8	14.6	29629	4	US-09-729-995-3
17	145.8	14.6	29629	4	US-10-135-689-3
18	145.6	14.5	866	4	US-09-257-179-11
19	145	14.5	1001	4	US-09-541-638-629
20	145	14.5	1001	4	US-10-170-097-629
21	145	14.5	20674	4	US-09-641-638-651
22	145	14.5	20674	4	US-10-170-097-651
23	144.4	14.4	1624	4	US-08-852-807-10
24	144.4	14.4	9837	1	US-08-832-883-68
25	144.4	14.4	9837	2	US-08-832-877-68
26	144.4	14.4	13674	2	US-08-852-807-1
27	144.2	14.4	137000	4	US-10-172-911-11

Query Match					15.5%	Score 154.8 ; DB 3 ; Length 1293 ;
Best Local Similarity					71.0%	Pred. No. 1.4e-32 ;
Matches 281 ;					Conservative 0 ;	Mismatches 22 ; Indels 93 ; Gaps 2 ;
Qy	107	CCAGGTCCAAATGCCTTGGAGACCTGGGACCCACCTATTTTAAAGCACTGGCGCTATGAAC	166			
Db	605	CTAAGTCCAAATGCCTTGGAGACCTGGGACCCACCTATTTTAAAGCACTGGCGCTATGAAC	664			
Qy	167	CACAAATTCAGCCCTACTGTCCCGTGTTCGCATTTGGGACCTGTGGCCAAAGCTGGAG	226			
Db	665	CACAAATTCAGCCCTACTGTCCCGTGTTCGCATTTGGGACCTGTGGCCAAAGCTGGAG	724			
Qy	227	GGACCTTCGAGGACCTGGCGTGTGTGGTGGTCCAAAGTGGGGGACGGTTCCTAGAGG	286			
Db	725	GGACCTTCGAGGACCTGGCGTGTGTGGTGGTCCAAAGTGGGGGACGGTTCCTAGAGG	749			
Qy	287	GCTCTGGGAGAGGTCCCGGCCCCACCCACCGGTGGAAGAAAGCTATGTCTATGTGACGGG	346			
Db	750	-----GGG	752			
Qy	347	TGGCTCTGTAGGCATCAGATTCACTGGATTGTGACCTGGACACCGGGACTCTGGCTG	406			
Db	753	TGGCTCTGTAGGCATCAGATTCACTGGATTGTGACCTGGACACCGGGACTCTGGCTG	812			
Qy	407	CTGGCCCTCACTACTCTTCCAGCTGCAGGAGAGAGCTACAACCTTCAGGTGAGGC-CCCA	465			
Db	813	CTGGCCCTCACTACTCTTCCAGCTGCAGGAGAGAGCTACAACCTTCAGGACGACTCA	872			
Qy	466	CTGTCCCACTGCTGCGCCCATCGCCCT	501			

ALIGNMENTS

RESULT 1  
US-09-381-681-1  
; Sequence 1, Application US/09381681  
; Patent No. 6255472  
; GENERAL INFORMATION:  
; APPLICANT: TAKIMURA, Takashi  
; APPLICANT: NAKAMURA, Yusuke  
; TITLE OF INVENTION: HUMAN GENES  
; FILE REFERENCE: Q55876  
; CURRENT APPLICATION NUMBER: US/09/381.681  
; EARLIER FILING DATE: 2000-01-10  
; EARLIER APPLICATION NUMBER: JPA 9-093044  
; EARLIER FILING DATE: 1997-03-26  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1293  
; TYPE: DNA  
; ORGANISM: Human  
US-09-381-681-1

C	28	144	14.4	15788	4	US-09-920-759-13	Sequence 13, Appl
	29	144	14.4	17493	4	US-09-804-471A-3	Sequence 3, Appl
	30	144	14.4	17493	4	US-10-238-709-3	Sequence 3, Appl
C	31	144	14.4	202001	4	US-09-734-674-3	Sequence 3, Appl
C	32	143.8	14.4	75395	4	US-09-884-890-3	Sequence 3, Appl
C	33	143.8	14.4	75395	4	US-10-274-194-3	Sequence 3, Appl
C	34	143.4	14.3	14636	3	US-09-173-914-6	Sequence 6, Appl
C	35	143.4	14.3	118999	4	US-09-791-105B-32	Sequence 32, Appl
C	36	143.4	14.3	169998	4	US-09-676-610B-24	Sequence 24, Appl
C	37	143.4	14.3	197496	4	US-09-877-177A-10	Sequence 10, Appl
C	38	143.4	14.3	197496	4	US-09-539-333D-1	Sequence 1, Appl
C	39	143.4	14.3	319608	4	US-09-879-409-1	Sequence 1, Appl
C	40	143.2	14.3	3001	4	US-09-539-333D-222	Sequence 222, App
C	41	142.8	14.3	31000	4	US-09-966-451-10	Sequence 10, Appl
C	42	142.8	14.3	55298	4	US-09-491-356C-1	Sequence 3, Appl
C	43	142.8	14.3	70000	4	US-09-851-896-3	Sequence 3, Appl
C	44	142.8	14.3	161652	4	US-09-497-855A-40	Sequence 40, Appl
C	45	142.6	14.2	36651	3	US-09-738-894A-3	Sequence 3, Appl



```

; ORGANISM: Homo sapiens;
US-09-497-855A-38

Query Match      15.3%; Score 153; DB 4; Length 128779;
Best Local Similarity 77.2%; Pred. No. 3.7e-31;
Matches 186; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 759 TTTTCATTTTACTTCCCTTCTCCCTTCAGCTTTGTTTTTTTTTTTAAAGACAGAT 818
Db 94246 TATAAATTATACATTTTAAATTTTTTATTTCTACTTTTTTTTTTTTTTTGGAGACAAGAT 94305

QY 819 CTCATTCTGTCAACCCAGGCTGGAGTGCAGTGGCCGACCTCGGCTCACTGTAACTCTGTC 878
Db 94306 CTTACTCTGTCAACCCAGGCTGGAGTGCATAGCGGATCTCGGCTCACTGCAACCTCTGC 94365

QY 879 TTCCTGGGTTCACCCAGTCTCTCTTCTCAGCTCTCTGAGTAGCTGGANTTACAGTGCT 938
Db 94366 CTCCTCAGGTTCAAGAGATTTCTTGCGCTCAGGCTCCCGAGTAGCTGGATTACAGGCACA 94425

QY 939 CGCCTACTCTCCAGCTAAATTTTATTTTGGTAGATAGAGATGGGTTTTTACAAATGTT 998
Db 94426 TGCACACAGCCCGAGCTAATTTTTTGTATTTTGGTAGACGAGTTTACCAATGCT 94485

QY 999 G 999
Db 94486 G 94486

```

```

RESULT 5
US-09-054-272-49/c
; Sequence 49, Application US/09054272
; Patent No. 6692909
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Daley, George O.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Rozen, Steven G.
; TITLE OF INVENTION: CODING SEQUENCE POLYMORPHISMS
; TITLE OF INVENTION: IN VASCULAR PATHOLOGY GENES
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/054,272
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH198-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; TELEX:
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13204 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-054-272-49

Query Match      15.0%; Score 150; DB 4; Length 13204;
Best Local Similarity 81.3%; Pred. No. 8.5e-31;
Matches 174; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 788 CAGCTTTGTTTTTTTTTTTTTTTAAAGACAGATCTCAATCTGTCAACCCAGGCTGGAGTGCAG 847
Db 7348 CTGCTTTTTTTTTTTTTTTTTTTTGGACAGAGTCTTGCTCTGTCAACCCAGGCTGGAGTGCAG 7289

QY 848 TGGCCCGACCTCGGCTCACTGTAACTCTGTCTCTGGGTTCAACCGAATTTCTCTTCTC 907
Db 7288 TGGCACACCTCTTGCTCACTGCAACCTCTGCTCTCTGGGTTCAAGTCAATTTCTCTGCTTC 7229

QY 908 AGCTCTCTGAGTAGCTGGAATTACAGTGCTCGGCACCTACTCCAGCTAATTTTATTT 967
Db 7228 AGCTCTCCGAGTAGCTGGAATTACAGGCATCGGTCAACACACCCGGCTAATTTTGTGT 7169

QY 968 TTGTTAGATAGAGATGGGTTTTTTCAAATGTTGGC 1001
Db 7168 ATTTTAGTAGAGATGGGTTTTTTCACATGTTGGC 7135

```

```

RESULT 6
US-09-851-896-3
; Sequence 3, Application US/09851896
; Patent No. 6410325
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI
; FILE REFERENCE: RTS-0220
; CURRENT APPLICATION NUMBER: US/09/851,896
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 70000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-851-896-3

```

```

Query Match      14.9%; Score 149.2; DB 4; Length 70000;
Best Local Similarity 81.9%; Pred. No. 3.1e-30;
Matches 172; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 792 TTTGTTTTTTTTTTTTTTTAAAGACAGAAATCTCAATCTGTCAACCCAGGCTGGAGTGCAGTGGC 851
Db 36071 TTTTTTTTTTTTTTTTTTGAGACGGAGTCTGCTCTGTGCCCCAGGCTGGAGTGCAGTGGC 36130

QY 852 CCGACCTCGGCTCACTGTAACTCTGTCTCTGGGTTCAACCGAATTTCTCTTCTCTCAGCC 911
Db 36131 CGGATCTCGGCTCACTGCAAGCTCCGCTCTCCGGGTTCAAGCCATTTCTCTGCTCAGCC 36190

QY 912 TCCTGAGTAGCTGGAATTACAGGTGCTCGGCACCTACTCCAGCTAATTTTATTTTGG 971
Db 36191 TCCGCGTAGCTGGGACTACAGGCGCCGCCACACACCCCGGCTAATTTTGTGTTTT 36250

QY 972 TAGATAGAGATGGGTTTTTTCAAATGTTGGC 1001
Db 36251 TTAGTAGAGACGGGGTTTTCACCATGTTGGC 36280

```

```

RESULT 7
US-09-749-588-3
; Sequence 3, Application US/09749588
; Patent No. 6423521
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al

```

```

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001068
; CURRENT APPLICATION NUMBER: US/09/749,588
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36159
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(36159)
; OTHER INFORMATION: n = A,T,C or G
US-09-749-588-3

```

```

Qy 827 GTACCCAGCGTGGAGTGCAGTGGCCGACCTTCGGCTACTGTAACTTCGTTCTCTGGG 886
Db 17986 ATTGCCAGGCTGGAGTGCAGTGGGGTAAATCTCAGCTCACTCAACCTCTGCCTCCCGG 18045
Qy 887 TTCAACGCGATTCTCTTCCTCCTCAGCCTCTCGAGTAGCTGGAATTACAGGTGCTCGCCACTA 946
Db 18046 TTCAAGCGATTCTCTCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGACGTTGCCACCA 18105
Qy 947 CTCGCCAGCTAAATTTTATATTTTGTAGATAGAGATGGGTTTTCACAATGTTGGC 1001
Db 18106 TGCCCGGCTAAATTTTGTATTTTTTTAGTAGACGGGGTTTTGCCATGTTGGC 18160

RESULT 9
US-09-916-204-3
; Sequence 3, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THERIOF
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 48763
; TYPE: DNA
; ORGANISM: Human
US-09-916-204-3

```

```

RESULT 8
US-10-135-687-3
; Sequence 3, Application US/10135687
; Patent No. 6689597
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01068DIV
; CURRENT APPLICATION NUMBER: US/10/135,687
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 09/749,588
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(36159)
; OTHER INFORMATION: n = A,T,C or G
US-10-135-687-3

Query Match 14.8%; Score 148.6; DB 4; Length 36159;
Best Local Similarity 77.0%; Fred. No. 3.3e-30;
Matches 181; Conservative 54; Indels 0; Gaps 0

QY 767 TTACTTCCCTCTCCCTTCAGCTTCTTTTAAAGACAGAATCTCATCT 826
|||
DB 17926 TTAGCTACTGTGTCCATTTTGTGAGACAGAGTCTCACTCT 17985
|||

```

```
; LENGTH: 48763
; TYPE: DNA
; ORGANISM: Human
US-10-282-048-3

Query Match      14.8%; Score 148.4; DB 4; Length 48763;
Best Local Similarity 73.7%; Pred. No. 4.3e-30;
Matches 202; Conservative 0; Mismatches 71; Indels 1; Gaps 1;

QY 728 TCTAGTATCTCCCTCCACAGCAATGGGGTGTTCATTTTACTTTCCCTTCTCCCTT 787
Db 16400 TTTAATCACTTCCCTAAAAGTCTACCTCGTTTTTTTTTTTGGTTTTTTT 16459

QY 788 CAGCTTTGTTTTTTTTTTTTTTAGACAGATCTCATTCTGCACCCAGGCTGAGTGCAG 847
Db 16460 TTTTITTTTTTTTTTTTTTTTGGTGTGTTAGTAGTCTTGCTCTGTCCACCCAGGCTGAGTGCAG 16519

QY 848 TGCCCGACCTCGGCTCACTTAACCTCTGCTTCTCTGGGTCAACCGATTCCTCTTCCTC 907
Db 16520 TGGTGGGATCTTGGCTCACTGCAACCTCCACCTCTCTGAGTCAAGCAATTCCTCGCTC 16579

QY 908 AGCTCTCTGAGTAGCTGGATTTACAGGTCTGCGCACTACTCCAGCTAATTTTATTT 967
Db 16580 AGCTCTCTGAGTAGCTGGGATTTATAGTGCCTGCCACACGCTGCTAATTTTGTGA 16639

QY 968 TTGTAGATAGAGATGGGTTTTTCACAATGTGGC 1001
Db 16640 TTTTAG-TAGAGTTGGGTTTTCACCATGTGGC 16672

RESULT 11
US-09-167-681-45
; Sequence 45, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinsilboum, M.D., Richard M.
; APPLICANT: Raitogianis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Otterness, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 847
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4361)...(4507)
; NAME/KEY: CDS
; LOCATION: (4612)...(4737)
; NAME/KEY: CDS
; LOCATION: (4827)...(4925)
; NAME/KEY: CDS
; LOCATION: (6322)...(6447)
; NAME/KEY: CDS
; LOCATION: (6543)...(6638)
; NAME/KEY: CDS
; LOCATION: (7137)...(7316)
; NAME/KEY: CDS
; LOCATION: (7439)...(7553)
; US-09-167-681-45

Query Match      14.7%; Score 147; DB 3; Length 8453;
Best Local Similarity 78.2%; Pred. No. 4.6e-30;
Matches 190; Conservative 0; Mismatches 50; Indels 3; Gaps 1;

QY 759 TTTCAATTTTACTTTCCCTTCCCTTCACCTTGTGTTTTTTTTTTTTTAAGACAGAA 818
Db 1668 TATCTCTCTGTTCTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1727

; US-09-491-356C-1
; Sequence 1, Application US/09491356C
; Patent No. 6566061
; GENERAL INFORMATION:
; APPLICANT: Philibert, Robert A.
; APPLICANT: Ginns, Edward I.
; APPLICANT: Delisi, Lynn
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13
; FILE REFERENCE: 9465.6US11
; CURRENT APPLICATION NUMBER: US/09/491,356C
; CURRENT FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: PCT/US99/09365
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 60/083,465
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 55298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (485)...(485)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (838)...(838)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (16728)...(16728)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (22750)...(22750)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (22756)...(22756)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (28519)...(28519)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (44804)...(44804)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (45002)...(45002)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (54049)...(54049)
; OTHER INFORMATION: n is not determined
; NAME/KEY: misc feature
; LOCATION: (54226)...(54226)
; OTHER INFORMATION: n is not determined
; US-09-491-356C-1
```

Query Match 14.7%; Score 147; DB 4; Length 55298;  
Best Local Similarity 81.0%; Pred. No. 1.1e-29;  
Matches 171; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
QY 791 CTTTGTCTTTTAAAGACAGAACTCTATTCTGTCACCCAGGCTGGAGTCCAGTGG 850  
Db 25227 CTTTGTCTTTTAAAGACAGAACTCTATTCTGTCACCCAGGCTGGAGTCCAGTGG 25286  
QY 851 CCGACCTCGCTCACTGTAACCTCTGCTCTGCTGCTCAACCGATCTCTCTCTCAGC 910  
Db 25287 CCGGATCTCGGCTCACTGTAACCTCTGCTCTGCTGCTCAACCGATCTCTCTCTCAGC 25346  
QY 911 CTCCTGAGTAGCTGAAATACAGGTCTCGGCACTACTCCAGCAATAATTTTATATTTG 970  
Db 25347 CTCCTGAGTAGCTGAAATACAGGTCTCGGCACTACTCCAGCAATAATTTTATATTTG 25406  
QY 971 GTAGATAGAGTGGTCTTTCACAAATGTTGGC 1001  
Db 25407 TTTGTAGACAGAGGTTCCACCATGTTGGC 25437

RESULT 13  
US-09-780-172-18/c  
; Sequence 18, Application US/09780172  
; Patent No. 6607916  
; GENERAL INFORMATION:  
; APPLICANT: Robert McKay  
; APPLICANT: Susan M. Freier  
; APPLICANT: Jacqueline Wyatt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION  
; FILE REFERENCE: R15-0159  
; CURRENT APPLICATION NUMBER: US/09/780,172  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 96  
; SEQ ID NO 18  
; LENGTH: 63000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-780-172-18

Query Match 14.7%; Score 147; DB 4; Length 63000;  
Best Local Similarity 81.6%; Pred. No. 1.2e-29;  
Matches 182; Conservative 0; Mismatches 40; Indels 1; Gaps 1;  
QY 779 TCTCCCTTCAGCTTTGTTTTTTTTTAAAGACAGAACTCTATTCTGTCCACCAAGCT 838  
Db 50904 TTTTCCCAATTTTTTTTTTTTTTTTTTTTGAAGACGAGTCTCGCTCTGTCCACCAAGCT 50845  
QY 839 GGAGTGCAGTCCCGACTCGCTCACTGTAACCTCTGCTTCTGGGTCAACCGATTG 898  
Db 50844 GGAGCAGATGCGCATCTCGGCTCACTGTAACCTCTCGCTCTCGGGTTCAAGCAATC 50785  
QY 899 TCTTCTCAGCTCTCGTAGTCTGGAATACAGGTCTCGCCACTACTCCAGCTAAT 958  
Db 50784 TCTTCTCAGCTCTCGTAGTCTGGAATACAGGTCTCGCCACTACTCCAGCTAAT 50725  
QY 959 TTTTATTTTGGTAGATAGATGGGTTTTCACAAATGTTGGC 1001  
Db 50724 TTTTATTTTGGTAGATAGATGGGTTTTCACAAATGTTGGC 50683

RESULT 14  
US-09-268-992-7/c  
; Sequence 7, Application US/09268992  
; Patent No. 6342351  
; GENERAL INFORMATION:  
; APPLICANT: Chen, H.  
; APPLICANT: Freimer, N.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
; FILE REFERENCE: 7853-138  
; CURRENT APPLICATION NUMBER: US/09/268,992

; CURRENT FILING DATE: 1999-03-16  
; EARLIER APPLICATION NUMBER: 09/236,134  
; EARLIER FILING DATE: 1999-01-22  
; EARLIER APPLICATION NUMBER: 60/106,056  
; EARLIER FILING DATE: 1998-10-28  
; EARLIER APPLICATION NUMBER: 60/088,312  
; EARLIER FILING DATE: 1998-06-05  
; EARLIER APPLICATION NUMBER: 60/078,044  
; EARLIER FILING DATE: 1998-03-16  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 72604  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: modified base  
; LOCATION: all n positions  
; OTHER INFORMATION: n=a, c, g, or t  
US-09-268-992-7

Query Match 14.6%; Score 146.2; DB 3; Length 72604;  
Best Local Similarity 83.3%; Pred. No. 2.1e-29;  
Matches 179; Conservative 0; Mismatches 33; Indels 3; Gaps 1;  
QY 789 AGCTTTGTTTTTTTTTAAAGACAGAACTCTATTCTGTCCACCAAGCTGAGTGCAGT 848  
Db 57752 AGCTTTGTTTTTTTTTAAAGACAGAACTCTATTCTGTCCACCAAGCTGAGTGCAGT 57693  
QY 849 GGGCCGACCTCGCTCACTGTAACCTCTCTTCTGGGTCAACCGATTCTCTTCTCTCA 908  
Db 57692 GGTGCGATCTCGCTCACTGTAACCTCTCTTCTGGGTCAACCGATTCTCTTCTCTCA 57633  
QY 909 GCCTCTGAGTAGCTGGAATACAGGTCTCGCCACTACTCCAGCTAATTTTATTTT 968  
Db 57632 GCCTCTGAGTAGCTGGAATACAGGTCTCGCCACTACTCCAGCTAATTTTATTTT 57573  
QY 969 ---TGCTAGATAGATGGGTTTTCACAAATGTTGG 1000  
Db 57572 GTATTTTAAAGATAGATGGGTTTTCACAAATGTTGG 57538

RESULT 15  
US-09-657-474-7/c  
; Sequence 7, Application US/09657474  
; Patent No. 6399762  
; GENERAL INFORMATION:  
; APPLICANT: Chen, H.  
; APPLICANT: Freimer, N.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
; FILE REFERENCE: 7853-138  
; CURRENT APPLICATION NUMBER: US/09/657,474  
; CURRENT FILING DATE: 2000-09-07  
; EARLIER APPLICATION NUMBER: 09/268,992  
; EARLIER FILING DATE: 1999-03-16  
; EARLIER APPLICATION NUMBER: 09/236,134  
; EARLIER FILING DATE: 1999-01-22  
; EARLIER APPLICATION NUMBER: 60/106,056  
; EARLIER FILING DATE: 1998-10-28  
; EARLIER APPLICATION NUMBER: 60/088,312  
; EARLIER FILING DATE: 1998-06-05  
; EARLIER APPLICATION NUMBER: 60/078,044  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 72604  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: modified base  
; LOCATION: all n positions



OTHER INFORMATION: n=a, c, g, or t  
US-09-657-474-7

Query Match	14.6%	Score	146.2	DB	3	Length	72604
Best Local Similarity	83.3%	Pred. No.	2.1e-29				
Matches	179	Conservative	0	Mismatches	33	Indels	3
				Gaps	1		

  

QY	789	AGCTTTGTTTTTTTTTTTTTTTAAAGACAGAACTCATTCGTACCCAGGCTGGAGTGCACT	848
Db	57752	AGCTTTTTTTTTTTTTTTTTTTTGGAGATGGAGTCTGGCTCTGTACCCAGGCTGGAGTGCAAT	57693
QY	849	GGCCCGACCTGGGCTCACTGTAACTTCCTTCCTGGGTCAACCGAATCTCCTTCCTCA	908
Db	57692	GGTGGATCTCGGCTCACTGCAACCTCTGCTCCTGGTTCAACAATCTCTGCTCA	57633
QY	909	GCTCTGTAGTAGCTGGATTACAGGTGTGCGCCACTACTCCAGCTAATTTTATATTT	968
Db	57632	GCTCTGTAGTAGCTGGATTACAGGCACGCGCCACCATGACCGCTAATATTTTTTTT	57573
QY	969	---TGCTAGATAGATGGGTTTTTCACAATGTTGG	1000
Db	57572	GTATTTTAAATAGATGGGGTTTCACCATGTTGG	57538

Search completed: November 21, 2004, 17:43:11  
Job time : 87.5046 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 479.756 Seconds  
(without alignments)  
10952.804 Million cell updates/sec

Title: US-09-820-095B-3\_COPY\_10000\_11000

Perfect score: 1001  
Sequence: 1 acacagtcaccccttcagcaag.....tgggttttcacaaatgttggc 1001

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	16449	8	AAL51049 Human P2X
2	393	39.3	576	4	AAL18586 Probe #85
3	393	39.3	576	4	AAL43700 Probe #12
4	393	39.3	576	4	ABA30783 Probe #92
5	393	39.3	576	4	AAL12112 Human bra
6	393	39.3	576	4	ABS37461 Human liv
7	393	39.3	576	6	ABS11822 Human gen
8	364	36.4	440	4	AAL12260 Probe #21
9	364	36.4	440	4	ABA53963 Human foe
10	364	36.4	440	4	AAL133615 Probe #23
11	364	36.4	440	4	ABA43513 Human bre
12	364	36.4	440	4	ABA23713 Probe #21
13	364	36.4	440	4	AAK227678 Human don
14	364	36.4	440	4	AAK02232 Human bra
15	364	36.4	440	4	ABS27249 Human liv
16	364	36.4	440	5	AAI02174 Probe #21
17	364	36.4	440	6	ABS02141 Human gen
18	161.2	16.1	9131	12	ADM97830 PXT-CMV-S
19	161.2	16.1	10557	12	ADM97832 PXT-CMV-S
20	157	15.7	7240	4	AAD04467 Human ins
21	157	15.7	7240	4	AAN31267 Human ins

C	22	157	15.7	7240	4	AAL31170 Human ins
C	23	157	15.7	7240	5	AAL50570 Insulin r
C	24	157	15.7	7240	6	ABQ72725 Human ins
C	25	157	15.7	7240	9	AAL62772 Human ins
C	26	155.6	15.5	63824	10	ADE43742 Human KNS
C	27	155.6	15.5	63824	10	ADE43743 Polymorph
C	28	155.6	15.5	63824	12	ADH54220 Human KNS
C	29	155.6	15.5	63824	12	ADH54221 Human KNS
C	30	155.6	15.5	153170	12	ADQ17382 Human sof
C	31	155.6	15.5	202100	10	ADE43315 Human IDE
C	32	155.6	15.5	202100	12	ADH54357 Human IDE
C	33	155.2	15.5	471	5	ABV61726 Human pro
C	34	154.8	15.5	1293	2	AAV61832 Coding se
C	35	154.8	15.5	1360	4	AAD04978 Human pur
C	36	154.8	15.5	1452	12	ADP49178 Human P2X
C	37	154.8	15.5	1452	12	ADP49172 Human P2X
C	38	154.8	15.5	1452	12	ADP49182 Human P2X
C	39	154.8	15.5	1452	12	ADP49174 Rat P2X2
C	40	154.8	15.5	1452	12	ADP49176 Human P2X
C	41	154.8	15.5	1697	2	AAV61833 Coding se
C	42	154.8	15.5	2693	8	AAL51048 Human P2X
C	43	153.8	15.4	14176	4	AAS26670 Human gen
C	44	153.8	15.4	14176	8	ABX74019 Human nov
C	45	153.8	15.4	177531	8	ACF62732 Cancer ba

## ALIGNMENTS

RESULT 1  
AAL51049  
ID AAL51049 standard; DNA; 16449 BP.  
XX  
AC AAL51049;  
XX  
DT 20-FEB-2003 (first entry)  
XX  
DE Human P2X-like purinergic receptor G-protein coupled receptor gene.  
XX  
KW Human; gene; ds; Gene therapy; G-protein coupled receptor; chromosome 22;  
KW P2X-like purinergic receptor; anaplastic oligodendroglioma; GPCR;  
KW chronic lymphocytic leukaemia; lung carcinoma; colon carcinoma;  
KW brain carcinoma; drug screening; tissue typing; pharmacogenomic analysis;  
KW SNP; single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT variation replace(136,C)  
FT /\*tag= a  
FT /\*note= "Single nucleotide polymorphism"  
FT variation replace(253,C)  
FT /\*tag= b  
FT /\*note= "Single nucleotide polymorphism"  
FT variation replace(573,T)  
FT /\*tag= c  
FT /\*note= "Single nucleotide polymorphism"  
FT variation replace(2000,G)  
FT /\*tag= d  
FT /\*note= "Single nucleotide polymorphism"  
FT CDS 2040..13451  
FT /\*tag= e  
FT /\*product= "Human P2X-like purinergic receptor-related G-protein coupled receptor"  
FT exon 2040..2095  
FT /\*tag= f  
FT /\*number= 1  
FT intron 2096..2776  
FT /\*tag= g  
FT /\*number= 1  
FT variation replace(222,C)  
FT /\*tag= h  
FT /\*note= "Single nucleotide polymorphism"



Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ACACAGTACCTTCAGCAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60
Db	10000 ACACAGTACCTTCAGCAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 10059
QY	61 CAAGACCTCTGTGTCCTTCTCATCTGACCTTCCACATCTCTCCAGGTCATATGCC 120
Db	10060 CAAGACCTCTGTGTCCTTCTCATCTGACCTTCCACATCTCTCCAGGTCATATGCC 10119
QY	121 TTGGAGACCTGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCC 180
Db	10120 TTGGAGACCTGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCC 10179
QY	181 TACTGTCCCGTGTTCGCAATTTGGGACCTCTGTGGCAAGCTGGAGGACCTTCAGGAC 240
Db	10180 TACTGTCCCGTGTTCGCAATTTGGGACCTCTGTGGCAAGCTGGAGGACCTTCAGGAC 10239
QY	241 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTCTCTAGAGGCTCTGGAGAGG 300
Db	10240 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTCTCTAGAGGCTCTGGAGAGG 10299
QY	301 TCCCGGGCCACCCACCGGTGGAAAGCTATGTCTATGTGAGGGTGGCTCTGTAGGCA 360
Db	10300 TCCCGGGCCACCCACCGGTGGAAAGCTATGTCTATGTGAGGGTGGCTCTGTAGGCA 10359
QY	361 TCAGATTCACTGGGATTGTGACCTGGACACCGGGGACTCTGGCTCTGGCTCTACT 420
Db	10360 TCAGATTCACTGGGATTGTGACCTGGACACCGGGGACTCTGGCTCTGGCTCTACT 10419
QY	421 CCTTCAGTGCAGGAGAGAGCTACAATTCAGGTGAGGCCCTCTCTCCAGTGC 480
Db	10420 CCTTCAGTGCAGGAGAGAGCTACAATTCAGGTGAGGCCCTCTCTCCAGTGC 10479
QY	481 AGTGTGGGCCCATCGCCCTCTCACTGTGGGGCCAGGACAGACACCCAGGCCAG 540
Db	10480 AGTGTGGGCCCATCGCCCTCTCACTGTGGGGCCAGGACAGACACCCAGGCCAG 10539
QY	541 GCTCTAGATATTCACCTAGCTGCAAGGGGTCCAGGAGCAGAGAGCTTCTC 600
Db	10540 GCTCTAGATATTCACCTAGCTGCAAGGGGTCCAGGAGCAGAGAGCTTCTC 10599
QY	601 AACCCACATCTCCAGCAGAGCTCGTCTCTGCCCCCAAGTCTGAGCCCTCCACCC 660
Db	10600 AACCCACATCTCCAGCAGAGCTCGTCTCTGCCCCCAAGTCTGAGCCCTCCACCC 10659
QY	661 CATCTGTCCAGGCCCTGCCCCAGCTCCTCACTGCGCCAGCCCTCTCCAGCCCA 720
Db	10660 CATCTGTCCAGGCCCTGCCCCAGCTCCTCACTGCGCCAGCCCTCTCCAGCCCA 10719
QY	721 CCTCGCTCTAGTATCTCCCTCCACAGCAATGGGTGTTTCAATTTTACTTTCCCTTC 780
Db	10720 CCTCGCTCTAGTATCTCCCTCCACAGCAATGGGTGTTTCAATTTTACTTTCCCTTC 10779
QY	781 TCCCTTTCAGCTTTGTTTTTTTTTTTTTTTAAAGACAGAACTCTATCTGTCACCAGGCTGG 840
Db	10780 TCCCTTTCAGCTTTGTTTTTTTTTTTTTTTAAAGACAGAACTCTATCTGTCACCAGGCTGG 10839
QY	841 AGTGCACTGCGGACCTCGGCTCACTGTAACCTCTGCTTCCGGTTCACCGATTCTC 900
Db	10840 AGTGCACTGCGGACCTCGGCTCACTGTAACCTCTGCTTCCGGTTCACCGATTCTC 10899
QY	901 CTTCCTCAGCTCTCTAGTAGCTGGAATTAACAGTGTCTGCGCACTACTCTCCAGCTAATTT 960
Db	10900 CTTCCTCAGCTCTCTAGTAGCTGGAATTAACAGTGTCTGCGCACTACTCTCCAGCTAATTT 10959
QY	961 TTATATTTTGGTAGATAGATGGTTTTTACAAATGTGGC 1001
Db	10960 TTATATTTTGGTAGATAGATGGTTTTTACAAATGTGGC 11000

RESULT 2  
AA118586/c

ID	AA118586 standard; DNA; 576 BP.
XX	AA118586;
AC	12-OCT-2001 (first entry)
DT	Probe #8519 for gene expression analysis in human cervical cell sample.
XX	Probe; human; microarray; gene expression; cervical epithelial cell;
DE	cervical cancer; ss.
XX	Homo sapiens.
OS	WO200157278-A2.
PN	09-AUG-2001.
XX	30-JAN-2001; 2001WO-US000670.
PF	04-FEB-2000; 2000US-0180312P.
XX	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.
PR	03-AUG-2000; 2000US-00632366.
PR	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.
PR	04-OCT-2000; 2000GB-00024263.
XX	(MOLE-) MOLECULAR DYNAMICS INC.
PA	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	WPI; 2001-489901/53.
XX	Human genome-derived single exon nucleic acid probes useful for analyzing
PT	gene expression in human cervical epithelial cells.
XX	Claim 25; SEQ ID NO 8519; 487pp; English.
XX	The present invention relates to human single exon nucleic acid probes
CC	(SENP). The present sequence is one such probe. The SENPs are derived
CC	from human HeLa cells. The SENPs can be used to produce a single exon
CC	microarray, which can be used for measuring human gene expression in a
CC	sample derived from human cervical epithelial cells. By measuring gene
CC	expression, the probes are therefore useful in grading and/or staging of
CC	diseases of the cervix, notably cervical cancer. Note: The sequence data
CC	for this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	Sequence 576 BP; 123 A; 158 C; 186 G; 109 T; 0 U; 0 Other;
SQ	Query Match 39.3%; Score 393; DB 4; Length 576;
	Best Local Similarity 100.0%; Pred. No. 5.2e-96;
	Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ACACAGTCACTTCAGCAAGTTCACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60
Db	393 ACACAGTCACTTCAGCAAGTTCACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 334
QY	61 CAAGACCTCTGTGTCCTTCTCATCTGACCTTCCACATCTCTCCAGGTCATATGCC 120
Db	333 CAAGACCTCTGTGTCCTTCTCATCTGACCTTCCACATCTCTCCAGGTCATATGCC 274
QY	121 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCC 180
Db	273 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCC 214
QY	181 TACTGTCCCGTGTTCGCAATTTGGGACCTCTGTGGCAAGCTGGAGGACCTTCAGGAC 240
Db	213 TACTGTCCCGTGTTCGCAATTTGGGACCTCTGTGGCAAGCTGGAGGACCTTCAGGAC 154
QY	241 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTCTCTAGAGGCTCTGGAGAGG 300

```

Db 153 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGGCAAGGTTCTAGAGGGTCTCGGAGAGGG 94
QY 301 TCCGGGGCCACCACCCGCTGGGAAAGCTATGTCTATGTGACAGGTTGCTCTGTAGGCA 360
Db 93 TCCGGGGCCACCACCCGCTGGGAAAGCTATGTCTATGTGACAGGTTGCTCTGTAGGCA 34
QY 361 TCAGAGTTCACTGGGATTTGACCTGGACACCG 393
Db 33 TCAGAGTTCACTGGGATTTGACCTGGACACCG 1

RESULT 3
AAI43700/c
ID AAI43700 standard; DNA; 576 BP.
XX AC AAI43700;
XX DT 17-OCT-2001 (first entry)
XX DE Probe #12386 used to measure gene expression in human placenta sample.
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
XX PS Claim 25; SEQ ID NO 12386; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders
XX SQ Sequence 576 BP; 123 A; 158 C; 186 G; 109 T; 0 U; 0 Other;

Query Match 39.3%; Score 393; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 5.2e-96;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTCACCTTCAGCAAGTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCC 60
Db 393 ACACAGTCACCTTCAGCAAGTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCC 334
QY 61 CAAGACCTCTCTGCTCCCTACCTCATCTGACCTTCCACTCTCCAGGTCATGCC 120
Db 333 CAAGACCTCTCTGCTCCCTACCTCATCTGACCTTCCACTCTCCAGGTCATGCC 274
QY 121 TTGAGAGCTGGGACCCCACTATTTTAAGCACTGCGCTATGAACCAATTCAGCC 180

```

```

Db 273 TTGAGAGCTGGGACCCCACTATTTTAAGCACTGCGCTATGAACCAATTCAGCC 214
QY 181 TACTGTCCCGTTTCCGCAATTGGGACCTCGTGGCCAAAGCTGAGGGACCTTCGAGGAC 240
Db 213 TACTGTCCCGTTTCCGCAATTGGGACCTCGTGGCCAAAGCTGAGGGACCTTCGAGGAC 154
QY 241 CTGCGGTTGCTGGTGGTCCCAAGTTGGGGGCAAGGTTCTAGAGGGCTCTGGAGAGGG 300
Db 153 CTGCGGTTGCTGGTGGTCCCAAGTTGGGGGCAAGGTTCTAGAGGGCTCTGGAGAGGG 94
QY 301 TCCGGGGCCACCACCCGCTGGGAAAGCTATGTCTATGTGACAGGTTGCTCTGTAGGCA 360
Db 93 TCCGGGGCCACCACCCGCTGGGAAAGCTATGTCTATGTGACAGGTTGCTCTGTAGGCA 34
QY 361 TCAGAGTTCACTGGGATTTGACCTGGACACCG 393
Db 33 TCAGAGTTCACTGGGATTTGACCTGGACACCG 1

RESULT 4
ABA30783/c
ID ABA30783 standard; DNA; 576 BP.
XX AC ABA30783;
XX DT 23-JAN-2002 (first entry)
XX DE Probe #9249 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX DR Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PS Claim 1; SEQ ID NO 9249; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from the human heart via microarrays. By measuring gene expression, the
XX CC probes are useful for predicting, diagnosing, grading, staging,
XX CC monitoring and prognosing diseases of the human heart and vascular system
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC congenital heart disease. Note: The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 576 BP; 123 A; 158 C; 186 G; 109 T; 0 U; 0 Other;

```

	Query Match	39.3%;	Score 393;	DB 4;	Length 576;
	Best Local Similarity	100.0%;	Pred. No. 5.2e-96;		
	Matches 393;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ACACAGTCACCTTCAGCAAGTTCAACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC	60		
Db	393	ACACAGTCACCTTCAGCAAGTTCAACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC	334		
Qy	61	CAAGACCTCTCTTGCCTACCTCATCTGACCTTCCCACTCCTCCAGGTCCAAATGCC	120		
Db	333	CAAGACCTCTCTTGCCTACCTCATCTGACCTTCCCACTCCTCCAGGTCCAAATGCC	274		
Qy	121	TTGAGAGACTGGGACCCCACTATTTAAGCACTGCCGCTATGAACCACAATTCAAGCCC	180		
Db	273	TTGAGAGACTGGGACCCCACTATTTAAGCACTGCCGCTATGAACCACAATTCAAGCCC	214		
Qy	181	TACTGTCCCGTTCGCGATTGGGGACCTCTGTGGCCAGGCTGGAGGACCTTCAGGAC	240		
Db	213	TACTGTCCCGTTCGCGATTGGGGACCTCTGTGGCCAGGCTGGAGGACCTTCAGGAC	154		
Qy	241	CTGCGTTGTGTTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGGCTCTGGGAGAGGG	300		
Db	153	CTGCGTTGTGTTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGGCTCTGGGAGAGGG	94		
Qy	301	TCCCGGGCCACCCACCGTGGAAAAGCTATGTGCTATGTGCAGGGTGGCTCTGTAGGCA	360		
Db	93	TCCCGGGCCACCCACCGTGGAAAAGCTATGTGCTATGTGCAGGGTGGCTCTGTAGGCA	34		
Qy	361	TCAGAGTTCACTGGGATTTGTGACTCTGGACACCG	393		
Db	33	TCAGAGTTCACTGGGATTTGTGACTCTGGACACCG	1		

RESULT 5	
AAK12112/c	
ID AAK12112 standard; DNA; 576 BP.	
XX	
AC AAK12112;	
XX	
DT 05-NOV-2001 (first entry)	
XX	
DE Human brain expressed single exon probe SEQ ID NO: 12103.	
XX	
KW Human; brain expressed exon; gene expression analysis; probe; microarray;	
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;	
XX ss.	
OS Homo sapiens.	
XX	
PN WC200157275-A2.	
XX	
PD 09-AUG-2001.	
XX	
PF 30-JAN-2001; 2001WO-US000667.	
XX	
PR 04-FEB-2000; 2000US-0180312P.	
PR	
PR 26-MAY-2000; 2000US-0207456P.	
PR	
PR 30-JUN-2000; 2000US-0060840B.	
PR	
PR 03-AUG-2000; 2000US-0063236E.	
PR	
PR 21-SEP-2000; 2000US-0234687P.	
PR	
PR 27-SEP-2000; 2000US-0236359P.	
PR	
PR 04-OCT-2000; 2000GB-00024263.	
XX	
PA (MOLE-) MOLECULAR DYNAMICS INC.	
XX	
XX	
PI Penn SG, Hanzel DK, Chen W, Rank DR;	
XX	
XX WPI; 2001-483446/52.	
DR	
XX	
XX	
PT Single exon nucleic acid probes for analyzing gene expression in human	
PT brains.	
XX	
XX	

```

PS Example 4; SEQ ID NO 12103; 650pp + Sequence Listing; English.
xx
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
xx
SQ Sequence 576 BP; 123 A; 158 C; 186 G; 109 T; 0 U; 0 Other;

Query Match 39.3%; Score 393; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 5.2e-36;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 ACACAGTCACCTTCAGCAAGTTCAACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC 60
DB 393 ACACAGTCACCTTCAGCAAGTTCAACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC 334

QY 61 CAAGACCGTCTTGTGCCCTACCTCATCTGACCTTTGCCACTCTCTCCAGGTCCAAATGCC 120
DB 333 CAAGACCGTCTTGTGCCCTACCTCATCTGACCTTTGCCACTCTCTCCAGGTCCAAATGCC 274

QY 121 TTGAGAGACCTGGGAGCCCCACCCTATTTAAAGCACCTGCCGCTATGAACACACAATTGACGCCCC 180
DB 273 TTGAGAGACCTGGGAGCCCCACCCTATTTAAAGCACCTGCCGCTATGAACACACAATTGACGCCCC 214

QY 181 TACTGTCCCGTGTCCGATTTGGGACCTCGTGGCCAAAGCTCGAGGAGACCTTCGAGGAC 240
DB 213 TACTGTCCCGTGTCCGATTTGGGACCTCGTGGCCAAAGCTCGAGGAGACCTTCGAGGAC 154

QY 241 CTGGCGTGTGCTGGTGGGTGCCAAGTCTGGGGCAGGGTTCCTTAGAGGGCTCTGGGAGAGGG 300
DB 153 CTGGCGTGTGCTGGTGGGTGCCAAGTCTGGGGCAGGGTTCCTTAGAGGGCTCTGGGAGAGGG 94

QY 301 TCCCGGGCCCAACCCACCGGTGGAAAGCTATGTCTATGTGCAGGGTGGCTCTGTAGGCA 360
DB 93 TCCCGGGCCCAACCCACCGGTGGAAAGCTATGTCTATGTGCAGGGTGGCTCTGTAGGCA 34

QY 361 TCAGAGTTCACTGGATTTGACCTGGACACCG 393
DB 33 TCAGAGTTCACTGGATTTGACCTGGACACCG 1

```

RESULT	6
ABS	37461/c
ID	ABS37461 standard; DNA; 576 BP.
XX	
AC	ABS37461;
XX	
DT	25-FEB-2003 (first entry)
XX	
DE	Human liver single exon probe, SEQ ID No 12451.
XX	
KW	Human; single exon nucleic acid probe; liver; cirrhosis;
KW	hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW	coronary heart disease; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200157273-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US000664.
XX	
PR	04-FEB-2000; 2000US-0180312P.
FR	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.
PR	03-AUG-2000; 2000US-00632366.
PR	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488898/53.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human adult liver.  
 PT Claim 1; SEQ ID NO 12451; 658pp; English.  
 PS  
 PS The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult liver.  
 CC (I) may be used for predicting, measuring and displaying gene expression  
 CC in samples derived from human adult liver. The genes identified may be  
 CC involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
 CC associated with coronary heart disease. ABS25011-ABS1005 represent human  
 CC liver single exon nucleic acid probes of the invention. Note: The  
 CC sequence information for this patent does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 576 BP; 123 A; 158 C; 186 G; 109 T; 0 U; 0 Other;  
 Query Match 39.3%; Score 393; DB 4; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-96;  
 Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCC 60  
 DB 393 ACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCC 334  
 QY 61 CAAGACCTTCCTGCTCCCTACTCTATCTGACCTTCCCTCCAGTCCCAATGCC 120  
 DB 333 CAAGACCTTCCTGCTCCCTACTCTATCTGACCTTCCCTCCAGTCCCAATGCC 274  
 QY 121 TTGGAGACTGGGACCCCACTTATTTAAGCACTGCCCTATGAACCAATTCAGCCCC 180  
 DB 273 TTGGAGACTGGGACCCCACTTATTTAAGCACTGCCCTATGAACCAATTCAGCCCC 214  
 QY 181 TACTCTCCGTTTCCGATTTGGGACCTCTGTCGACCTGTCGAGGACCTTCGAGGAC 240  
 DB 213 TACTCTCCGTTTCCGATTTGGGACCTCTGTCGACCTGTCGAGGACCTTCGAGGAC 154  
 QY 241 CTGGGCTTCTGCTGGGTGCTCCCAAGTTGGGGGAGGGTTCCTAGAGGGCTCTGGGAGAGG 300  
 DB 153 CTGGGCTTCTGCTGGGTGCTCCCAAGTTGGGGGAGGGTTCCTAGAGGGCTCTGGGAGAGG 94  
 QY 301 TCCCGGGCCACCCACCGTGGAAGCTATGTCTATGTCAGGCTGCTCTAGGCA 360  
 DB 93 TCCCGGGCCACCCACCGTGGAAGCTATGTCTATGTCAGGCTGCTCTAGGCA 34  
 QY 361 TCAGAGTTCACTGGGATTTGACCTGGGACCG 393  
 DB 33 TCAGAGTTCACTGGGATTTGACCTGGGACCG 1  
 RESULT 7  
 ABS11822/c  
 ID ABS11822 standard; DNA; 576 BP.  
 XX  
 AC ABS11822;  
 XX  
 XX 19-AUG-2002 (first entry)  
 XX Human genome-derived single exon probe from lung SEQ ID No 11813.  
 DE  
 XX

KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200186003-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000665.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2002-114183/15.  
 DR  
 XX Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples.  
 PT  
 XX Claim 1; SEQ ID NO 11813; 634pp; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 16614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC; the novel set of probes which hybridise at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a sample  
 CC derived from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of the  
 CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberculous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a single exon probe of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed



CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 576 BP; 123 A; 158 C; 186 G; 109 T; 0 U; 0 Other;  
Query Match 39.3%; Score 393; DB 6; Length 576;  
Best Local Similarity 100.0%; Pred. No. 5.2e-96;  
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACACAGTCACTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60  
Db 393 ACACAGTCACTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 334  
QY 61 CAAGACCCCTTCCTGTCCTCCCTACTCATCTGACCTTTCCCACTCCCTCCAGGTCCCAATGCC 120  
Db 333 CAAGACCCCTTCCTGTCCTCCCTACTCATCTGACCTTTCCCACTCCCTCCAGGTCCCAATGCC 274  
QY 121 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180  
Db 273 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 214  
QY 181 TACTGTCCCGTGTTCGGCATTTGGGACCTCGTGGCCNAGGCTGGAGGACCTTCGAGGAC 240  
Db 213 TACTGTCCCGTGTTCGGCATTTGGGACCTCGTGGCCNAGGCTGGAGGACCTTCGAGGAC 154  
QY 241 CTGGCGTTCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTAGAGGGTCTCTGGGAGAGGG 300  
Db 153 CTGGCGTTCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTAGAGGGTCTCTGGGAGAGGG 94  
QY 301 TCCCGGGCCACCCAGCGGTGAAAGCTATGTGCTATGTGAGGGTGGCTCTGTAGGCA 360  
Db 93 TCCCGGGCCACCCAGCGGTGAAAGCTATGTGCTATGTGAGGGTGGCTCTGTAGGCA 34  
QY 361 TCAGAGTTCACCTGGGATTTGACCTGGACACCG 393  
Db 33 TCAGAGTTCACCTGGGATTTGACCTGGACACCG 1  
RESULT 8  
AA112260/c  
ID AA112260 standard; DNA; 440 BP.  
XX  
AC AA112260;  
XX  
XX  
DT 12-OCT-2001 (first entry)  
DE Probe #2193 for gene expression analysis in human cervical cell sample.  
DE Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer; ss.  
XX Homo sapiens.  
XX  
XX WO200157278-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US0000670.  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
XX  
XX 26-MAY-2000; 2000US-0207456P.  
XX  
XX 30-JUN-2000; 2000US-00608408.  
XX  
XX 03-AUG-2000; 2000US-00632366.  
XX  
XX 21-SEP-2000; 2000US-0234687P.  
XX  
XX 27-SEP-2000; 2000US-0236359P.  
XX  
XX 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488901/53.  
XX  
XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human cervical epithelial cells.  
XX  
XX Claim 25; SEQ ID NO 2193; 487pp; English.  
XX  
CC The present invention relates to human single exon nucleic acid probes  
CC (SENPs). The present sequence is one such probe. The SENPs are derived  
CC from human HeLa cells. The SENPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging of  
CC diseases of the cervix, notably cervical cancer. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 440 BP; 96 A; 115 C; 143 G; 86 T; 0 U; 0 Other;  
SQ  
Query Match 36.4%; Score 364; DB 4; Length 440;  
Best Local Similarity 100.0%; Pred. No. 3.3e-88;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACACAGTCACTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60  
Db 364 ACACAGTCACTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 305  
QY 61 CAAGACCCCTTCCTGTCCTCCCTACTCATCTGACCTTTCCCACTCCCTCCAGGTCCCAATGCC 120  
Db 304 CAAGACCCCTTCCTGTCCTCCCTACTCATCTGACCTTTCCCACTCCCTCCAGGTCCCAATGCC 245  
QY 121 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180  
Db 244 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 185  
QY 181 TACTGTCCCGTGTTCGGCATTTGGGACCTCGTGGCCNAGGCTGGAGGACCTTCGAGGAC 240  
Db 184 TACTGTCCCGTGTTCGGCATTTGGGACCTCGTGGCCNAGGCTGGAGGACCTTCGAGGAC 125  
QY 241 CTGGCGTTCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTAGAGGGTGGCTCTGGGAGAGGG 300  
Db 124 CTGGCGTTCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTAGAGGGTGGCTCTGGGAGAGGG 65  
QY 301 TCCCGGGCCACCCAGCGGTGAAAGCTATGTGCTATGTGAGGGTGGCTCTGTAGGCA 360  
Db 64 TCCCGGGCCACCCAGCGGTGAAAGCTATGTGCTATGTGAGGGTGGCTCTGTAGGCA 5  
QY 361 TCAG 364  
Db 4 TCAG 1  
RESULT 9  
ABA53963/c  
ID ABA53963 standard; DNA; 440 BP.  
XX  
XX ABA53963;  
XX  
XX 01-FEB-2002 (first entry)  
DT  
DE Human foetal liver single exon nucleic acid probe #2268.  
DE Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
XX Homo sapiens.  
XX  
XX WO200157277-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US0000669.  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
XX  
XX 26-MAY-2000; 2000US-0207456P.  
XX

PR	30-JUN-2000; 2000US-00608408.
PR	03-AUG-2000; 2000US-00632366.
PR	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.
PR	04-OCT-2000; 2000GB-00024263.
XX	(MOLE-) MOLECULAR DYNAMICS INC.
PA	
XX	Penn SG, Hanzel DK, Chen W, Rank DR;
PI	WPI; 2001-483447/52.
XX	
XX	Human genome-derived single exon nucleic acid probes useful for analyzing
PT	gene expression in human fetal liver.
XX	
XX	Claim 1; SEQ ID NO 2268; 639pp + Sequence Listing; English.
XX	
CC	The invention relates to a single exon nucleic acid probe for measuring
CC	human gene expression in a sample derived from human fetal liver. The
CC	single exon nucleic acid probes may be used for predicting, measuring and
CC	displaying gene expression in samples derived from human fetal liver. The
CC	present sequence is a single exon nucleic acid probe of the invention.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at <a href="http://ftp.wipo.int/pub/published_pct_sequences">ftp.wipo.int/pub/published_pct_sequences</a>
XX	
XX	Sequence 440 BP; 96 A; 115 C; 143 G; 86 T; 0 U; 0 Other;
XX	
XX	Query Match 36.4%; Score 364; DB 4; Length 440;
XX	Best Local Similarity 100.0%; Pred.No. 3.3e-88;
XX	Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCC 60
DB	364 ACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCC 305
QY	61 CAAGACCTCTGTGCCCTACCTCATCTGACCTTTCCCACTCTCCAGGTCCAATGCC 120
DB	304 CAAGACCTCTGTGCCCTACCTCATCTGACCTTTCCCACTCTCCAGGTCCAATGCC 245
QY	121 TTGGAGACCTGGGACCCACCACTATTTAAGCACTGGCGCTATGAACCAATTCAGCCCC 180
DB	244 TTGGAGACCTGGGACCCACCACTATTTAAGCACTGGCGCTATGAACCAATTCAGCCCC 185
QY	181 TACTGTCCCGTGTTCGGCATTTGGGGACCTCGTGGCCAAAGCTCGAGGACCTTCGAGGAC 240
DB	184 TACTGTCCCGTGTTCGGCATTTGGGGACCTCGTGGCCAAAGCTCGAGGACCTTCGAGGAC 125
QY	241 CTGGGGTTGCTGGTGGGTCCCAAGTTGGGGGCGAGGTTCTTAGAGGGCTCTGGGAGAGGG 300
DB	124 CTGGGGTTGCTGGTGGGTCCCAAGTTGGGGGCGAGGTTCTTAGAGGGCTCTGGGAGAGGG 65
QY	301 TCCCGGGCCCCACCCACCGCTGCGAAAGCTATGTGCTATGTGACAGGTTGCTCTGTAGGCA 360
DB	64 TCCCGGGCCCCACCCACCGTGGAAAGCTATGTGCTATGTGACAGGTTGCTCTGTAGGCA 5
QY	361 TCAG 364
DB	4 TCAG 1
XX	
XX	RESULT 10
ID	AAI33615/c
ID	AAI33615 standard; DNA; 440 BP.
XX	
XX	AAI33615;
XX	
XX	17-OCT-2001 (first entry)
DE	Probe #2301 used to measure gene expression in human placenta sample.
XX	
KW	Probe; microarray; human; placenta; antenatal diagnosis;
KW	genetic disorder; ss.

```
AC ABA43513;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human breast cell single exon nucleic acid probe #2208.
XX
XX Human; microarray; single exon probe; gene expression; breast; disease;
XX cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157271-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000662.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
XX for measuring gene expression in sample derived from human breast,
XX comprises number of single exon nucleic acid probes.
XX
XX Claim 1; SEQ ID NO 2208; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting the
XX probes with a collection of detectably labelled nucleic acids derived
XX from mRNA of human breast, and then measuring the label bound to each
XX probe of the microarray. The probes are useful for verifying the
XX expression of regions of genomic DNA predicted to encode proteins. They
XX are useful for gene discovery, and for determining predisposition and/or
XX assessing the toxicity of chemical agents on cells. The microarray of
XX this invention presents a far greater diversity of probes for measuring
XX gene expression, with far less bias than expressed sequence tag
XX microarrays. The method is suitable for rapid production of functional
XX information from genomic sequence. The present sequence is a single exon
XX nucleic acid probe of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 440 BP; 96 A; 115 C; 143 G; 86 T; 0 U; 0 Other;
XX
XX Query Match 36.4%; Score 364; DB 4; Length 440;
XX Best Local Similarity 100.0%; Pred. No. 3.3e-88;
XX Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ACACAGTCACCTTCAGCAAGTTCACCTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60
XX
XX 364 ACACAGTCACCTTCAGCAAGTTCACCTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 305
XX
XX 61 CAAGACCTCTTGTCCCTTACTCTACTCTTCTTCACTTCTTCCACTCTCCAGGTCCTAATGCC 120
XX
XX 304 CAAGACCTCTTGTCCCTTACTCTACTCTTCTTCACTTCTTCCACTCTCCAGGTCCTAATGCC 245
XX
XX 121 TTGGAGACCTGGACCCCTTATTTTAAAGCACTGGCGTATGACCAAAATTCAGCCCC 180
XX
XX 244 TTGGAGACCTGGACCCCTTATTTTAAAGCACTGGCGTATGACCAAAATTCAGCCCC 185
```

```
QY 181 TACTGTCCCGTGTTCGCGATTGGGACCTCGTGCCCAAGGCTGGAGGACCTTCGAGGAC 240
DB 184 TACTGTCCCGTGTTCGCGATTGGGACCTCGTGCCCAAGGCTGGAGGACCTTCGAGGAC 125
QY 241 CTGGCGTTGCTGGTGGGTCCCAAGTTGGGGCAGGGTCTCTAGAGGGCTCTGGGAGAGGG 300
DB 124 CTGGCGTTGCTGGTGGGTCCCAAGTTGGGGCAGGGTCTCTAGAGGGCTCTGGGAGAGGG 65
QY 301 TCCCGGGCCACCCACCGGTGTAAGGCTATGTCTATGTGCAGGGTGGCTCTCTAGGCA 360
DB 64 TCCCGGGCCACCCACCGGTGTAAGGCTATGTCTATGTGCAGGGTGGCTCTCTAGGCA 5
QY 361 TCAG 364
DB 4 TCAG 1
```

```
RESULT 12
ABA23713/c
ID ABA23713 standard; DNA; 440 BP.
XX
XX ABA23713;
XX
XX 23-JAN-2002 (first entry)
XX
XX Probe #2179 for gene expression analysis in human heart cell sample.
XX
XX Human; gene expression; heart; microarray; vascular system; probe;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX
XX Claim 1; SEQ ID NO 2179; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart. The
XX present sequence is one such probe. The probes may be used for
XX predicting, measuring and displaying gene expression in samples derived
XX from the human heart via microarrays. By measuring gene expression, the
XX probes are useful for predicting, diagnosing, grading, staging,
XX monitoring and prognosing diseases of the human heart and vascular system
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX congenital heart disease. Note: The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 440 BP; 96 A; 115 C; 143 G; 86 T; 0 U; 0 Other;
```

Query Match 36.4%; Score 364; DB 4; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-88;  
 Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTCCCTTACGAAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60  
 DB ACACAGTCCCTTACGAAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 305

QY 61 CAAGACCTCTTGTCCCTTACCTCATCTGACCTTTCCCACTCTCCAGGTCCAATGCC 120  
 DB CAAGACCTCTTGTCCCTTACCTCATCTGACCTTTCCCACTCTCCAGGTCCAATGCC 245

QY 121 TTGGAGACCTGGGACCCCACTTATTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180  
 DB TTGGAGACCTGGGACCCCACTTATTTAAGCACTGCGCTATGAACCAATTCAGCCCC 185

QY 181 TACTGTCCCGTGTTCGCGCATTTGGGACCTTCGTGGCCAAAGCTATGTGCTATGTGAGGCA 240  
 DB TACTGTCCCGTGTTCGCGCATTTGGGACCTTCGTGGCCAAAGCTATGTGCTATGTGAGGCA 125

QY 241 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGGCTCTGGGAGAGG 300  
 DB CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGGCTCTGGGAGAGG 65

QY 301 TCCCGGGCCACCCACCGGTGGAAGCTATGTGCTATGTGAGGCA 360  
 DB TCCCGGGCCACCCACCGGTGGAAGCTATGTGCTATGTGAGGCA 5

QY 361 TCAG 364  
 DB 4 TCAG 1

## RESULT 13

AAK27678/c  
 ID AAK27678 standard; DNA; 440 BP.

XX AAK27678;  
 AC  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human bone marrow expressed single exon probe SEQ ID NO: 2235.

XX Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000669.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-0060840P.  
 PR 03-AUG-2000; 2000US-0063236P.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX  
 DR WPI; 2001-488900/53.  
 XX  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.  
 XX  
 XX Example 4; SEQ ID NO 2235; 658pp + Sequence Listing; English.  
 PS  
 XX

CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of  
 CC the probes of the invention.

SQ Sequence 440 BP; 96 A; 115 C; 143 G; 86 T; 0 U; 0 Other;

Query Match 36.4%; Score 364; DB 4; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-88;  
 Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTCCCTTACGAAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60  
 DB 364 ACACAGTCCCTTACGAAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 305

QY 61 CAAGACCTCTTGTCCCTTACCTCATCTGACCTTTCCCACTCTCCAGGTCCAATGCC 120  
 DB 304 CAAGACCTCTTGTCCCTTACCTCATCTGACCTTTCCCACTCTCCAGGTCCAATGCC 245

QY 121 TTGGAGACCTGGGACCCCACTTATTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180  
 DB 244 TTGGAGACCTGGGACCCCACTTATTTAAGCACTGCGCTATGAACCAATTCAGCCCC 185

QY 181 TACTGTCCCGTGTTCGCGCATTTGGGACCTTCGTGGCCAAAGCTATGTGCTATGTGAGGCA 240  
 DB 184 TACTGTCCCGTGTTCGCGCATTTGGGACCTTCGTGGCCAAAGCTATGTGCTATGTGAGGCA 125

QY 241 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGGCTCTGGGAGAGG 300  
 DB 124 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGGCTCTGGGAGAGG 65

QY 301 TCCCGGGCCACCCACCGGTGGAAGCTATGTGCTATGTGAGGCA 360  
 DB 64 TCCCGGGCCACCCACCGGTGGAAGCTATGTGCTATGTGAGGCA 5

QY 361 TCAG 364  
 DB 4 TCAG 1

## RESULT 14

AAK02232/c  
 ID AAK02232 standard; DNA; 440 BP.

XX AAK02232;  
 AC  
 XX  
 DT 05-NOV-2001 (first entry)  
 DE Human brain expressed single exon probe SEQ ID NO: 2223.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;  
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
 KW ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-0060840P.  
 PR 03-AUG-2000; 2000US-0063236P.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.

```
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 2223; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention
XX
XX Sequence 440 BP; 96 A; 115 C; 143 G; 86 T; 0 U; 0 Other;
XX
XX Query Match 36.4%; Score 364; DB 4; Length 440;
XX Best Local Similarity 100.0%; Pred. No. 3.3e-88;
XX Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ACACAGTCCCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60
XX 364 ACACAGTCCCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 305
XX
XX 61 CAAGACCTCTTGTCCCTACCTCATCTGACCTTTCCACCTCTCCCTCCAGGTCGAATGCC 120
XX 304 CAAGACCTCTTGTCCCTACCTCATCTGACCTTTCCACCTCTCCCTCCAGGTCGAATGCC 245
XX
XX 121 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180
XX 244 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 185
XX
XX 181 TACTGTCCCGTGTCCGCAATTTGGGACCTCGTGCCCAAGGCTGGAGGACCTTCGAGGAC 240
XX 184 TACTGTCCCGTGTCCGCAATTTGGGACCTCGTGCCCAAGGCTGGAGGACCTTCGAGGAC 125
XX
XX 241 CTGGCGTTCTGTGGTGGTCCCAAGTTGGGGCAGGGTTCTTAGAGGGCTCTGGGAGGG 300
XX 124 CTGGCGTTCTGTGGTGGTCCCAAGTTGGGGCAGGGTTCTTAGAGGGCTCTGGGAGGG 65
XX
XX 301 TCCCGGGCCACCCACCGGTGGAAGACTATGTGCTATGTGAGGGTGGCTCTGTAGGCA 360
XX 64 TCCCGGGCCACCCACCGGTGGAAGACTATGTGCTATGTGAGGGTGGCTCTGTAGGCA 5
XX
XX 361 TCAG 364
XX 4 TCAG 1
XX
XX RESULT 15
XX ABS27249/C
XX ID ABS27249 standard; DNA; 440 BP.
XX AC ABS27249;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver single exon probe, SEQ ID NO 2239.
XX
XX Human; single exon nucleic acid probe; liver; cirrhosis;
XX KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX KW coronary heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
```

```
PF 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 1; SEQ ID NO 2239; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENp) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABS25011-ABS51005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 440 BP; 96 A; 115 C; 143 G; 86 T; 0 U; 0 Other;
XX
XX Query Match 36.4%; Score 364; DB 4; Length 440;
XX Best Local Similarity 100.0%; Pred. No. 3.3e-88;
XX Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ACACAGTCCCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60
XX 364 ACACAGTCCCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 305
XX
XX 61 CAAGACCTCTTGTCCCTACCTCATCTGACCTTTCCACCTCTCCCTCCAGGTCGAATGCC 120
XX 304 CAAGACCTCTTGTCCCTACCTCATCTGACCTTTCCACCTCTCCCTCCAGGTCGAATGCC 245
XX
XX 121 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180
XX 244 TTGGAGACCTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 185
XX
XX 181 TACTGTCCCGTGTCCGCAATTTGGGACCTCGTGCCCAAGGCTGGAGGACCTTCGAGGAC 240
XX 184 TACTGTCCCGTGTCCGCAATTTGGGACCTCGTGCCCAAGGCTGGAGGACCTTCGAGGAC 125
XX
XX 241 CTGGCGTTCTGTGGTGGTCCCAAGTTGGGGCAGGGTTCTTAGAGGGCTCTGGGAGGG 300
XX 124 CTGGCGTTCTGTGGTGGTCCCAAGTTGGGGCAGGGTTCTTAGAGGGCTCTGGGAGGG 65
XX
XX 301 TCCCGGGCCACCCACCGGTGGAAGACTATGTGCTATGTGAGGGTGGCTCTGTAGGCA 360
XX 64 TCCCGGGCCACCCACCGGTGGAAGACTATGTGCTATGTGAGGGTGGCTCTGTAGGCA 5
XX
XX 361 TCAG 364
XX 4 TCAG 1
XX
XX Search completed: November 21, 2004, 18:28:58
XX Job time : 483.756 secs
```

Sun Nov 28 09:38:17 2004

us-09-820-095b-3\_copy\_10000\_11000.rng

Page 12

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 4283.34 Seconds  
(without alignments)  
11051.408 Million cell updates/sec

Title: US-09-820-095B-3\_COPY\_10000\_11000  
Perfect score: 1001  
Sequence: 1 acacagtcaccttcgcaag.....tgggttttcacaaatgtggc 1001

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	98.9	147086	9	AC002472 Homo sapi
2	990	98.9	162470	9	AC007664 Homo sapi
3	949.8	94.9	28984	9	AB002059 Homo sapi
4	827.4	82.7	175167	2	AC116044 Papio ham
5	333	39.3	576	6	CQ072719 Sequence
6	333	39.3	576	6	CQ103527 Sequence
7	333	39.3	576	6	CQ177853 Sequence
8	333	39.3	576	6	CQ225612 Sequence
9	333	39.3	576	6	CQ300708 Sequence
10	333	39.3	576	6	CQ338009 Sequence
11	364	36.4	440	6	CQ051345 Sequence
12	364	36.4	440	6	CQ066393 Sequence
13	364	36.4	440	6	CQ093442 Sequence
14	364	36.4	440	6	CQ132213 Sequence
15	364	36.4	440	6	CQ170783 Sequence
16	364	36.4	440	6	CQ199931 Sequence
17	364	36.4	440	6	CQ215400 Sequence
18	364	36.4	440	6	CQ254007 Sequence
19	364	36.4	440	6	CQ291027 Sequence

C	20	364	36.4	440	6	CQ328129 Sequence
	21	229.8	23.0	203973	10	AC115733 Mus muscu
	22	218.8	21.9	164055	2	AC130883 Rattus no
C	23	162.6	16.2	176184	9	AC018845 Homo sapi
	24	162.6	16.2	185281	9	AC007338 Homo sapi
C	25	162	16.2	180750	2	AC097327 Pan trogl
	26	161.2	16.1	127431	9	AC027796 Homo sapi
	27	161.2	16.1	159397	2	AC027040 Homo sapi
C	28	161.2	16.1	200237	9	AF168787 Homo sapi
C	29	158.8	15.9	157069	9	AC010163 Homo sapi
	30	158.8	15.9	161506	9	AF450105 Human DNA
C	31	158.8	15.9	169250	9	AF002812 Homo sapi
	32	158.6	15.8	53645	2	AC084246 Homo sapi
C	33	158.2	15.8	64425	9	AL591368 Human DNA
	34	158.2	15.8	72172	9	AC010311 Homo sapi
	35	158.2	15.8	315681	2	AL353380 Homo sapi
	36	157.8	15.8	147285	2	AC148838 Pan trogl
	37	157.6	15.7	125990	9	AC108734 Homo sapi
	38	157.4	15.7	64923	9	AC002545 Homo sapi
	39	157.4	15.7	65386	2	AC002532 Homo sapi
C	40	157.4	15.7	100634	9	AP001594 Homo sapi
	41	157.4	15.7	340000	9	AP001695 Homo sapi
C	42	157.2	15.7	149320	2	AC126357 Homo sapi
C	43	157.2	15.7	153053	9	AC007533 Homo sapi
C	44	157.2	15.7	182101	9	AC007599 Homo sapi
C	45	157	15.7	7240	6	AX114613 Sequence

ALIGNMENTS

RESULT 1  
AC002472  
LOCUS Homo sapiens Chromosome 22q11.2 PAC Clone p\_n5 In BCRL2-GGT Region, complete sequence.  
ACCESSION AC002472  
VERSION AC002472.8 GI:24137490  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 147086)  
AUTHORS Budarf,M.L. and Emanuel,B.S.  
JOURNAL Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 2 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Lao,V. and Roe,B.A.  
JOURNAL Homo sapiens Chromosome 22q11.2 PAC Clone p\_n5 In BCRL2-GGT Region  
REFERENCE 3 (bases 1 to 147086)  
AUTHORS Zhang,G., Lao,V., Zhan,M. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (20-AUG-1997) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA  
REFERENCE 4 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (10-SEP-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA  
REFERENCE 5 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (23-OCT-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA  
REFERENCE 6 (bases 1 to 147086)  
AUTHORS Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.  
JOURNAL Direct Submission  
TITLE Submitted (13-NOV-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 7 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (18-NOV-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	ORIGIN	Query Match Best Local Similarity 98.9%; Score 990; DB 9; Length 147086; Matches 1001; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 8 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (21-NOV-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 1	ACACAGTCACTTCAGCAAGTTCACAACTTCCTTAAGTAAGCAGAGTGGTCTCATCTGCC 60
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 9 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (14-APR-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67314	ACACAGTCACTTCAGCAAGTTCACAACTTCCTTAAGTAAGCAGAGTGGTCTCATCTGCC 67313
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 10 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (13-SEP-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 61	CAAGACCTTCCTGTCCTCCCTACCTCATCTGACCTTCCCACTCCCTCCAGGTCCTCAATGCC 120
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 11 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (13-SEP-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67374	CAAGACCTTCCTGTCCTCCCTACCTCATCTGACCTTCCCACTCCCTCCAGGTCCTCAATGCC 67433
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 12 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 121	TTGGAGACCTTGGGACCCCACTTATTTTAAGCACTCCCGCTATGAACCAATTCAGCCCC 180
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 13 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (20-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67434	TTGGAGACCTTGGGACCCCACTTATTTTAAGCACTCCCGCTATGAACCAATTCAGCCCC 67493
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 14 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (31-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 181	TACTGTCCGTCCTGTCCTCCGATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCGGAGAC 240
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 15 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67494	TACTGTCCGTCCTGTCCTCCGATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCGGAGAC 67553
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 16 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 241	CTGGCGTTCGTGTCCTGTCCTCCGATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCGGAGAC 300
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 17 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67554	CTGGCGTTCGTGTCCTGTCCTCCGATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCGGAGAC 67613
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 18 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 301	TCCCGGCGCCACCCACCGTGGAAAGCTATGTCATGTCAGGGTGGCTCTGTAGGCA 360
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 19 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67614	TCCCGGCGCCACCCACCGTGGAAAGCTATGTCATGTCAGGGTGGCTCTGTAGGCA 67673
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 20 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 361	TCAGATTTCAGTGGGATTCGACCTGGACACCGGGGACCTCTGGCTCTGTCGCTCTACT 420
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 21 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67674	TCAGATTTCAGTGGGATTCGACCTGGACACCGGGGACCTCTGGCTCTGTCGCTCTACT 67733
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 22 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 421	CCTTCCAGTTCAGGAGAGAGCTACAACTTCAGTGGGCGCCCACTGCTCCAGTGGCC 480
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 23 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67734	CCTTCCAGTTCAGGAGAGAGCTACAACTTCAGTGGGCGCCCACTGCTCCAGTGGCC 67793
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 24 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 481	AGTGTGTGGGCGCCATCGCCCTCTCACTGTGGCGGCGAGACACACACCCAGCCAG 540
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 25 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67794	AGTGTGTGGGCGCCATCGCCCTCTCACTGTGGCGGCGAGACACACACCCAGCCAG 67853
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 26 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 541	GCCTCTAGATATTCACCTACGTGTGCAAGGGGTCCAGGAGCAGGAGAGCTCTCTC 600
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 27 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67854	GCCTCTAGATATTCACCTACGTGTGCAAGGGGTCCAGGAGCAGGAGAGCTCTCTC 67913
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 28 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 601	AACCCACATTCCTCAGCAGACAGCTCCGTCCTGTCGCCCCCAAGCTCTGAGCCCTCCACCC 560
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 29 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67914	AACCCACATTCCTCAGCAGACAGCTCCGTCCTGTCGCCCCCAAGCTCTGAGCCCTCCACCC 67973
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 30 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 661	CATCTGTCCAGGCGCCCTGCCAGCTCAGCTCCCTCACTGCCAGCCCTTCCTCCACCCCA 720
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 31 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 67974	CATCTGTCCAGGCGCCCTGCCAGCTCAGCTCCCTCACTGCCAGCCCTTCCTCCACCCCA 68033
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 32 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 721	CCTCGCTTCCTAGTATCTCCCTCCACAGCAATGGGGTGTTCATTTTACTTTTCCCTTC 780
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 33 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 68034	CCTCGCTTCCTAGTATCTCCCTCCACAGCAATGGGGTGTTCATTTTACTTTTCCCTTC 68093
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 34 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 781	TCCCTTCAGCTTTG-TTTTTTTTTTTTTTTTAAAGACAGAACTCTCATTTCTGCACCCAGGCTG 839
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 35 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 68094	TCCCTTCAGCTTTG-TTTTTTTTTTTTTTTTAAAGACAGAACTCTCATTTCTGCACCCAGGCTG 68153
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 36 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 840	GAGTGCAGTGGGCGGACCTCGGCTCACTGTAACCTCTGCTTCTGGTTCACACGATTC 899
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 37 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 68154	GAGTGCAGTGGGCGGACCTCGGCTCACTGTAACCTCTGCTTCTGGTTCACACGATTC 68213
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 38 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Qy 900	CCTTCTTCAGCTTCCTGAGTACTGGAATTACAGGTGTCGCGACCTACTCCAGCTAAT 959
REFERENCE AUTHORS TITLE JOURNAL	OK 73019, USA 39 (bases 1 to 147086) Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A. Direct Submission Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	Db 68214	CCTTCTTCAGCTTCCTGAGTACTGGAATTACAGGTGTCGCGACCTACTCCAGCTAAT 68273

COMMENT  
On Oct 19, 2002 this sequence version replaced gi:22597497.  
Because these overlapping clones came from different libraries.

FEATURES  
source  
1..147086  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"







QY 899 TCCTTCTCAGCTCCTGAGTACTGGAATTACAGGTGCTGCGCACTATCCAGCTAAT 958  
 Db 17964 TCCTTCTCAGCTCCTGAGTACTGGAATTACAGGTGCTGCGCACTATCCAGCTAAT 18023  
 QY 959 TTTTATATTGTTAGATAGATAGATGGTTTTCACATGTTGGC 1001  
 Db 18024 TTTTATATTGTTG---TAGAGATGGTTTTCACATGTTGGC 18062

RESULT 4  
 AC116044  
 LOCUS  
 DEFINITION  
 Papi0 hamadryas clone RP41-7017, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 11  
 ordered pieces.  
 AC116044  
 VERSION  
 HTG; HTGS\_PHASE2; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 KEYWORDS  
 SOURCE  
 Papi0 hamadryas (hamadryas baboon)  
 ORGANISM  
 Papi0 hamadryas  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopitheciinae; Papio.  
 1 (bases 1 to 175167)  
 Birren, B., Nusbaum, C. and Lander, E.  
 Papi0 hamadryas, clone RP41-7017  
 Unpublished  
 2 (bases 1 to 175167)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
 Boukhgalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J.,  
 Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
 Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,  
 Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
 MacLean, C., MacDonald, P., Major, J., Marguis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,  
 Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
 Submitted (23-MAR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 175167)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
 Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,  
 Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,  
 Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K.,  
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,  
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,  
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B.,  
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C.,  
 MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,  
 Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,  
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,  
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 COMMENT

Direct Submission  
 Submitted (04-FEB-2004) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Feb 4, 2004 this sequence version replaced gi:19657483.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 -----  
 Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WtBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu  
 -----  
 Project Information  
 Center project name: L12529  
 Center clone name: 7\_O\_17

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 11 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* been provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

\* 1 6404: contig of 6404 bp in length  
 \* 6405 6504: gap of 100 bp  
 \* 6505 12196: contig of 5692 bp in length  
 \* 12197 12297: gap of 100 bp  
 \* 12297 15159: contig of 2863 bp in length  
 \* 15160 15259: gap of 100 bp  
 \* 15260 23315: contig of 10056 bp in length  
 \* 23316 25415: gap of 100 bp  
 \* 25416 25416: contig of 5399 bp in length  
 \* 25417 79513: gap of 100 bp  
 \* 79514 79514: contig of 10292 bp in length  
 \* 79514 89805: gap of 100 bp  
 \* 89806 92601: contig of 2696 bp in length  
 \* 92602 92701: gap of 100 bp  
 \* 92702 101172: contig of 8471 bp in length  
 \* 101173 101272: gap of 100 bp  
 \* 101273 114938: contig of 13666 bp in length  
 \* 114939 115038: gap of 100 bp  
 \* 115039 141113: contig of 26075 bp in length  
 \* 141114 141213: gap of 100 bp  
 \* 141214 175167: contig of 33954 bp in length.

FEATURES  
 source

1. 175167  
 /organism="Papi0 hamadryas"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9557"  
 /clone="RP41-7017"  
 /clone\_lib="RPCI-41 Male Baboon BAC"

ORIGIN

Query Match 82.7% Score 827.4; DB 2; Length 175167;  
 Best Local Similarity 91.6%; Pred. No. 1.9e-205;  
 Matches 924; Conservative 0; Mismatches 71; Indels 14; Gaps 4;  
 Qy 1 ACACAGTCACCTTCAGCAAGTTCACCTCTCTAAGTAAGCAGAGTGGTCTCATCTGCCC 60  
 Db 153698 ACACGGTCACCTTCAGCAAGTTCACCTCTCTAAGTAAGCAGAGTGGTCTCATCTGCCC 153747  
 Qy 61 CAAGACCTCTCTGTGCTCCCTTACCTCTATCTGACCTTTCCCACTCTCCAGGTCCAATGCC 120  
 Db 153748 CAAGACCTCTCTGTGCTCTCTTACCTCTGCTGACCTCTCCCACTCCCGCCAGGTCCAATGCC 153807  
 Qy 121 TTGAGACCTTGGGACCCCACTATTATTAAGCACTGGCGGTATGAACCAACAATTCAGCCCC 180  
 Db 153808 TTGAGACCTTGGGACCCCACTATTATTAAGCACTGGCGGTATGAACCAACAATTCAGCCCC 153867  
 Qy 181 TACTGTCCGTGTTCGCGATTGGGACCTCGTGGCCAGCGCTCGAGGACCTTCGAGGAC 240

```
Db 153868 TACTGTCCGCTTCCGCAATGGGACCTCGTGCAGGCTGGAGGAACTTCAGGAC 153927
QY 241 CTGGGCTTCTGGTGGGTCCTCAAGTTGGGGCAGGGTTCCTAGAGGGCTCTGGAGAGGG 300
Db 153928 CTGGGCTTCTGGTGGGTCCTCAAGTTGGGGCAGGGTTCCTAGAGGGCTCTGGAGAGGG 153987
QY 301 TCCGGGGCCACCCACCGTGGAAAGCTATGTGCTATGTGCAGGGTGGCTCTTAGGCA 360
Db 153988 TCCGGGGCCACCCACCGTGGAAAGCTATGTGCTATGTGCAGGGTGGCTCTTAGGCA 154047
QY 361 TCAGAGTTCACTGGGATTTGACCTGGACACGGGGACTCTGGCTGTGGCTCACTACT 420
Db 154048 TCAGAGTTCACTGGGATTTGACCTGGGACTCTGGGACTCTGGGCTGTGGCTCACTACT 154107
QY 421 CCTTCAGCTGAGGAGAGAGCTACAACTTCAGGTGAGGCCCCCAGCTCTCCAGTCCCC 480
Db 154108 CCTTCAGCTGAGGAGAGAGCTACAACTTCAGGTGAGGCCCCCAGCTCTCCAGTCCCC 154167
QY 481 ACCTGCTGGGCCATTCGCCCTCTCACTGTGGGGCCAGGACAGACACACCGAGGCCAG 540
Db 154168 ACCTGCTGGGCCATTCGCCCTCTCACTGTGGGGCCAGGACAGACACACCGAGGCCAG 154227
QY 541 GCCTCTAGATATTCACCTACGTTGCAAGGGGTCCTCAGGAGCAGGAGAGAGCTGTCTC 600
Db 154228 CCCTCTGATATTCACCT- GGTTTCAGAGGGGCTCCAGAGCAGGAGAGAGCTGTCTC 154286
QY 601 AACCCACATCTCCAGACAGGCTCCGTCCTGTGTCGCCCAAGTCTGAGCCCTCCACCC 660
Db 154287 AACCCACATCTCCAGACAGGCTCCATCTGTCTCCGCCCAAGTCTGAGCCCTCCACCT 154346
QY 661 CATCTGTCCAGGCCCTGCCAGCTCAGGCTCCTCACTGCAGCCCTCTCTCCACCCCA 720
Db 154347 CATCTGTCCAGGCCCTGCCAGCTCAGGCTCCTCACTGCAGCCCTCTCTCCACCCCA 154406
QY 721 CCTCGCTTCTAGTATCTCCCTCCACAGCAATGGGGTGTTCATTTTACCTTTCCCTTC 780
Db 154407 CCTCGCTTCTAGTATCTCCCTCCACAGCAATGGGGTGTTCATTTTACCTTTCCCTTC 154466
QY 781 TCCCTTCTAGCTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 832
Db 154467 TCCGTTTCTAGCTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 154526
QY 833 CAGGTGGAGTGCAGTGGCCGACCTCGCTCACTGTAACTCTGCTTCTCGGTTCACAC 892
Db 154527 CAGGTGGAGTGCAGTGGCCGACCTCTGCTCACTGTAACTCTCGCTTCCGGGGTCAAC 154586
QY 893 CGATTCTCTCTCAGCTCTCTGAGTGTGGAATTCAGAGTGTCTCCCACTACTCCCA 952
Db 154587 CGATTCTCTCTCAGCTCTCTGAGTGTGGAATTCAGAGTGTCTCCCACTACTCCCA 154645
QY 953 GCTAATTTTATTTTGTAGATAGATGGTTTTTACCAATGTTCG 1001
Db 154646 GCTAATTTTATTTTGG-----TAGAGATGGTTTTTACCAATGTTCG 154690

RESULT 5
LOCUS CQ072719/c
DEFINITION Sequence 8519 from Patent WO0157278.
ACCESSION CQ072719
VERSION CQ072719.1 GI:41042588
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human hela cells or other human cervical epithelial cells
JOURNAL Patent: WO 0157278-A 8519 09-AUG-2001;
Aeomica, Inc. (US)
```

```
FEATURES
source
Location/Qualifiers
1..576
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002472.3-EXPRESSED IN HELA, SIGNAL = 1.4"

ORIGIN
Query Match 39.3%; Score 393; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 6.6e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTCACTTCAGCAGGTTCACTCTCTTAAGTAAAGCAGAGTGGTCTCATCTGCC 60
Db 393 ACACAGTCACTTCAGCAGGTTCACTCTCTTAAGTAAAGCAGAGTGGTCTCATCTGCC 334
QY 61 CAAGACCTCTCTTGTCCCTTACCTCACTGACCTTTCCCACTCTCCAGGTCCAATGCC 120
Db 333 CAAGACCTCTCTTGTCCCTTACCTCACTGACCTTTCCCACTCTCCAGGTCCAATGCC 274
QY 121 TTGGAGACTTGGGACCCCACTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180
Db 273 TTGGAGACTTGGGACCCCACTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 214
QY 181 TACTGTCCCGTGTTCGGCAATTGGGACCTCTGCGCAAGGCTGGAGGACCTTCGAGGAC 240
Db 213 TACTGTCCCGTGTTCGGCAATTGGGACCTCTGCGCAAGGCTGGAGGACCTTCGAGGAC 154
QY 241 CTGGCGTTCTGCTGGTGGTCCCAAGTTGGGGCAGGGTCTTAGAGGGCTCTGGAGAGGG 300
Db 153 CTGGCGTTCTGCTGGTGGTCCCAAGTTGGGGCAGGGTCTTAGAGGGCTCTGGAGAGGG 94
QY 301 TCCCGGGCCACCCACCGTGGAAAGCTATGTGTCAGAGGTGGCTCTGTAGGCA 360
Db 93 TCCCGGGCCACCCACCGTGGAAAGCTATGTGTCAGAGGTGGCTCTGTAGGCA 34
QY 361 TCAGAGTTCACTGGGATTTGACCTGGACACCG 393
Db 33 TCAGAGTTCACTGGGATTTGACCTGGACACCG 1

RESULT 6
LOCUS CQ103527/c
DEFINITION Sequence 12386 from Patent WO0157272.
ACCESSION CQ103527
VERSION CQ103527.1 GI:41072578
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human placenta
JOURNAL Patent: WO 0157272-A 12386 09-AUG-2001;
Aeomica, Inc. (US)

FEATURES
source
Location/Qualifiers
1..576
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002472.3-EXPRESSED IN PLACENTA, SIGNAL = 1.6"

ORIGIN
Query Match 39.3%; Score 393; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 6.6e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTCACTTCAGCAGGTTCACTCTCTTAAGTAAAGCAGAGTGGTCTCATCTGCC 60
```

Db 393 ACACAGTCACCTTACAGAAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 334  
Qy 61 CAGAGCCCTCTTGTCTCCCTACTCTACTCTGACCTTTCACCTTCCCAAGTCCCAATGCC 120  
Db 333 CAGAGCCCTCTTGTCTCCCTACTCTACTCTGACCTTTCACCTTCCCAAGTCCCAATGCC 274  
Qy 121 TTGGAGACCTTGGAGCCCACTTATTTTAAGCACTGCCGCTATGAACCAAAATTCAGCCCC 180  
Db 273 TTGGAGACCTTGGAGCCCACTTATTTTAAGCACTGCCGCTATGAACCAAAATTCAGCCCC 214  
Qy 181 TACTGTCCCGTGTTCGGATTTGGGACCTCTGTCGCAAGCTGAGGAGGAGGACCTTCGAGGAC 240  
Db 213 TACTGTCCCGTGTTCGGATTTGGGACCTCTGTCGCAAGCTGAGGAGGAGGACCTTCGAGGAC 154  
Qy 241 CTGGCGTTCCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTCTAGAGGGCTCTGGAGAGGG 300  
Db 153 CTGGCGTTCCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTCTAGAGGGCTCTGGAGAGGG 94  
Qy 301 TCCCGGGCCACCCACCGTGGAAAGCTATGTCTATGTGCAAGGTTGGTCTGTAGGCA 360  
Db 93 TCCCGGGCCACCCACCGTGGAAAGCTATGTCTATGTGCAAGGTTGGTCTGTAGGCA 34  
Qy 361 TCAGAGTTCACCTGGGATTCGACCTGGACACCG 393  
Db 33 TCAGAGTTCACCTGGGATTCGACCTGGACACCG 1  
RESULT 7  
LOCUS CQ177853/c 576 bp DNA linear PAT 21-JAN-2004  
DEFINITION Sequence 9249 from Patent WO0157274.  
ACCESSION CQ177853  
VERSION CQ177853.1 GI:41172592  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human heart  
JOURNAL Patent: WO 0157274-A 9249 09-AUG-2001;  
FEATURES  
source location/Qualifiers  
1. 576  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AC002472.3-EXPRESSED IN HEART, SIGNAL = 4.1"  
ORIGIN  
Query Match 39.3%; Score 393; DB 6; Length 576;  
Best Local Similarity 100.0%; Pred. No. 6.6e-92;  
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACACAGTCACCTTACAGAAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60  
Db 393 ACACAGTCACCTTACAGAAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 334  
Qy 61 CAGAGCCCTCTTGTCTCCCTACTCTACTCTGACCTTTCACCTTCCCAAGTCCCAATGCC 120  
Db 273 TTGGAGACCTTGGAGCCCACTTATTTTAAGCACTGCCGCTATGAACCAAAATTCAGCCCC 214  
Qy 181 TACTGTCCCGTGTTCGGATTTGGGACCTCTGTCGCAAGCTGAGGAGGAGGACCTTCGAGGAC 240  
Db 213 TACTGTCCCGTGTTCGGATTTGGGACCTCTGTCGCAAGCTGAGGAGGAGGACCTTCGAGGAC 154  
Qy 241 CTGGCGTTCCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTCTAGAGGGCTCTGGAGAGGG 300  
Db 153 CTGGCGTTCCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTCTAGAGGGCTCTGGAGAGGG 94

Db 153 CTGGCGTTCCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTCTAGAGGGCTCTGGAGAGGG 94  
Qy 301 TCCCGGGCCACCCACCGTGGAAAGCTATGTCTATGTGCAAGGTTGGTCTGTAGGCA 360  
Db 93 TCCCGGGCCACCCACCGTGGAAAGCTATGTCTATGTGCAAGGTTGGTCTGTAGGCA 34  
Qy 361 TCAGAGTTCACCTGGGATTCGACCTGGACACCG 393  
Db 33 TCAGAGTTCACCTGGGATTCGACCTGGACACCG 1  
RESULT 8  
LOCUS CQ225612/c 576 bp DNA linear PAT 21-JAN-2004  
DEFINITION Sequence 12451 from Patent WO0157273.  
ACCESSION CQ225612  
VERSION CQ225612.1 GI:41208499  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO 3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US 60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03 August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000 (03.10.00)<150> US 60/234,359<151> 27 September 2000 (27.09.00)<150> US 60/236,687<151> 21 September 2000 (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>  
JOURNAL Molecular Dynamics Sequence Listing Engine  
FEATURES  
source location/Qualifiers  
1. 576  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AC002472.3-EXPRESSED IN ADULT LIVER, SIGNAL = 1.9"  
ORIGIN  
Query Match 39.3%; Score 393; DB 6; Length 576;  
Best Local Similarity 100.0%; Pred. No. 6.6e-92;  
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACACAGTCACCTTACAGAAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 60  
Db 393 ACACAGTCACCTTACAGAAAGTTCAACTTCTTAAGTAAGCAGAGTGGGTCTCATCTGCC 334  
Qy 61 CAGAGCCCTCTTGTCTCCCTACTCTACTCTGACCTTTCACCTTCCCAAGTCCCAATGCC 120  
Db 333 CAGAGCCCTCTTGTCTCCCTACTCTACTCTGACCTTTCACCTTCCCAAGTCCCAATGCC 274  
Qy 121 TTGGAGACCTTGGAGCCCACTTATTTTAAGCACTGCCGCTATGAACCAAAATTCAGCCCC 180  
Db 273 TTGGAGACCTTGGAGCCCACTTATTTTAAGCACTGCCGCTATGAACCAAAATTCAGCCCC 214  
Qy 181 TACTGTCCCGTGTTCGGATTTGGGACCTCTGTCGCAAGCTGAGGAGGAGGACCTTCGAGGAC 240  
Db 213 TACTGTCCCGTGTTCGGATTTGGGACCTCTGTCGCAAGCTGAGGAGGAGGACCTTCGAGGAC 154  
Qy 241 CTGGCGTTCCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTCTAGAGGGCTCTGGAGAGGG 300  
Db 153 CTGGCGTTCCTGTGGGTCCCAAGTTGGGGGCGAGGTTCTTCTAGAGGGCTCTGGAGAGGG 94  
Qy 301 TCCCGGGCCACCCACCGTGGAAAGCTATGTCTATGTGCAAGGTTGGTCTGTAGGCA 360  
Db 93 TCCCGGGCCACCCACCGTGGAAAGCTATGTCTATGTGCAAGGTTGGTCTGTAGGCA 34

```
QY 361 TCAGAGTTCACCTGGGATTGTGACCTGGACACCG 393
Db 33 TCAGAGTTCACCTGGGATTGTGACCTGGACACCG 1

RESULT 9
LOCUS CQ300708 576 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 11813 from Patent WO0186003.
ACCESSION CQ300708
VERSION CQ300708.1 GI:41261285
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human lung
JOURNAL Patent: WO 0186003-A 11813 15-NOV-2001;
Aeomica, Inc. (US)
FEATURES
source Location/Qualifiers
1..576
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002472.3-EXPRESSED IN LUNG, SIGNAL = 1.5"

ORIGIN
Query Match 39.3%; Score 393; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 6.6e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTTCACCTTCAGCAAGTTCACCTTCTCTAAGTAGCAGAGTGGGTCTCATCTGCC 60
Db 393 ACACAGTTCACCTTCAGCAAGTTCACCTTCTCTAAGTAGCAGAGTGGGTCTCATCTGCC 334

QY 61 CAAGACCTTCCTTGTCCCTACCTCATCTGACCTTCCCACTCCCTCCAGGTCCAATGCC 120
Db 333 CAAGACCTTCCTTGTCCCTACCTCATCTGACCTTCCCACTCCCTCCAGGTCCAATGCC 274

QY 121 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTCCGCTATGAACCAATTCAGCCCC 180
Db 273 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTCCGCTATGAACCAATTCAGCCCC 214

QY 181 TACTGTCCCGTGTTCGGCATTCGGGACCTCTGTGGCCAAAGGCTTGGAGGACCTTCGAGGAC 240
Db 213 TACTGTCCCGTGTTCGGCATTCGGGACCTCTGTGGCCAAAGGCTTGGAGGACCTTCGAGGAC 154

QY 241 CTGGCGTTCTGGTGGTCCCAAGTTGGGGCAGAGTTCCTAGAGGGCTCTGGGAGAGGG 300
Db 153 CTGGCGTTCTGGTGGTCCCAAGTTGGGGCAGAGTTCCTAGAGGGCTCTGGGAGAGGG 94

QY 301 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGTCAGGGTGGCTCTGTAGGCA 360
Db 93 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGTCAGGGTGGCTCTGTAGGCA 34

QY 361 TCAGAGTTCACCTGGGATTGTGACCTGGACACCG 393
Db 33 TCAGAGTTCACCTGGGATTGTGACCTGGACACCG 1

RESULT 11
LOCUS CQ051345/c 440 bp DNA linear PAT 19-JAN-2004
DEFINITION Sequence 2165 from Patent WO0157270.
ACCESSION CQ051345
VERSION CQ051345.1 GI:41025851
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human breast and hbl 100 cells
JOURNAL Patent: WO 0157270-A 2165 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source Location/Qualifiers
1..440
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002472.3-EXPRESSED IN HBL100, SIGNAL = 5"

ORIGIN
Query Match 39.3%; Score 393; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 6.6e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTTCACCTTCAGCAAGTTCACCTTCTCTAAGTAGCAGAGTGGGTCTCATCTGCC 60
Db 393 ACACAGTTCACCTTCAGCAAGTTCACCTTCTCTAAGTAGCAGAGTGGGTCTCATCTGCC 334

QY 61 CAAGACCTTCCTTGTCCCTACCTCATCTGACCTTCCCACTCCCTCCAGGTCCAATGCC 120
Db 333 CAAGACCTTCCTTGTCCCTACCTCATCTGACCTTCCCACTCCCTCCAGGTCCAATGCC 274

QY 121 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTCCGCTATGAACCAATTCAGCCCC 180
Db 273 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTCCGCTATGAACCAATTCAGCCCC 214

QY 181 TACTGTCCCGTGTTCGGCATTCGGGACCTCTGTGGCCAAAGGCTTGGAGGACCTTCGAGGAC 240
Db 213 TACTGTCCCGTGTTCGGCATTCGGGACCTCTGTGGCCAAAGGCTTGGAGGACCTTCGAGGAC 154

QY 241 CTGGCGTTCTGGTGGTCCCAAGTTGGGGCAGAGTTCCTAGAGGGCTCTGGGAGAGGG 300
Db 153 CTGGCGTTCTGGTGGTCCCAAGTTGGGGCAGAGTTCCTAGAGGGCTCTGGGAGAGGG 94

QY 301 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGTCAGGGTGGCTCTGTAGGCA 360
Db 93 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGTCAGGGTGGCTCTGTAGGCA 34

QY 361 TCAGAGTTCACCTGGGATTGTGACCTGGACACCG 393
Db 33 TCAGAGTTCACCTGGGATTGTGACCTGGACACCG 1

RESULT 10
LOCUS CQ338009/c 576 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 12103 from Patent WO0157275.
ACCESSION CQ338009
VERSION CQ338009.1 GI:41287080
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
AUTHORS Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human breast and hbl 100 cells
JOURNAL Patent: WO 0157270-A 2165 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source Location/Qualifiers
1..440
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002472.3-EXPRESSED IN HBL100, SIGNAL = 5"

ORIGIN
Query Match 39.3%; Score 393; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 6.6e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTTCACCTTCAGCAAGTTCACCTTCTCTAAGTAGCAGAGTGGGTCTCATCTGCC 60
Db 393 ACACAGTTCACCTTCAGCAAGTTCACCTTCTCTAAGTAGCAGAGTGGGTCTCATCTGCC 334

QY 61 CAAGACCTTCCTTGTCCCTACCTCATCTGACCTTCCCACTCCCTCCAGGTCCAATGCC 120
Db 333 CAAGACCTTCCTTGTCCCTACCTCATCTGACCTTCCCACTCCCTCCAGGTCCAATGCC 274

QY 121 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTCCGCTATGAACCAATTCAGCCCC 180
Db 273 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTCCGCTATGAACCAATTCAGCCCC 214

QY 181 TACTGTCCCGTGTTCGGCATTCGGGACCTCTGTGGCCAAAGGCTTGGAGGACCTTCGAGGAC 240
Db 213 TACTGTCCCGTGTTCGGCATTCGGGACCTCTGTGGCCAAAGGCTTGGAGGACCTTCGAGGAC 154

QY 241 CTGGCGTTCTGGTGGTCCCAAGTTGGGGCAGAGTTCCTAGAGGGCTCTGGGAGAGGG 300
Db 153 CTGGCGTTCTGGTGGTCCCAAGTTGGGGCAGAGTTCCTAGAGGGCTCTGGGAGAGGG 94

QY 301 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGTCAGGGTGGCTCTGTAGGCA 360
Db 93 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGTCAGGGTGGCTCTGTAGGCA 34

QY 361 TCAGAGTTCACCTGGGATTGTGACCTGGACACCG 393
Db 33 TCAGAGTTCACCTGGGATTGTGACCTGGACACCG 1
```

```
Query Match      36.4%; Score 364; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.5e-84;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTCACCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCC 60
Db 364 ACACAGTCACCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCC 305

QY 61 CAAGACCCCTTCCTGTCCTCCCTACCTCATCTGACCTTTCCACCTCCCTCCAGGTCCAATGCC 120
Db 304 CAAGACCCCTTCCTGTCCTCCCTACCTCATCTGACCTTTCCACCTCCCTCCAGGTCCAATGCC 245

QY 121 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCCGTATGAACCAATTCAGCCCC 180
Db 244 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCCGTATGAACCAATTCAGCCCC 185

QY 181 TACTGTCCCGTGTTCGGCAATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCAGGAC 240
Db 184 TACTGTCCCGTGTTCGGCAATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCAGGAC 125

QY 241 CTGGCGTTCTGCTGGTCCCAAGTTGGGGCAGAGTTCCTTAGAGGGTCTTGGGAGAGGG 300
Db 124 CTGGCGTTCTGCTGGTCCCAAGTTGGGGCAGAGTTCCTTAGAGGGTCTTGGGAGAGGG 65

QY 301 TCCGGGGCCACCCACCGGTGAAAAGCTATGTCTATGTGACAGGTGGTCTGTAGGCA 360
Db 64 TCCGGGGCCACCCACCGGTGAAAAGCTATGTCTATGTGACAGGTGGTCTGTAGGCA 5

QY 361 TCAG 364
Db 4 TCAG 1

RESULT 12
CQ066393/c
LOCUS      440 bp      DNA      linear      PAT 20-JAN-2004
DEFINITION Sequence 2193 from Patent WO0157278.
ACCESSION CQ066393
VERSION    CQ066393.1 GI:41036262
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelial cells
JOURNAL   Patent: WO 0157278-A 2193 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES   Location/Qualifiers
            source
            1..440
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="MAP TO AC002472.3-EXPRESSED IN HELA, SIGNAL = 4.9"

ORIGIN
Query Match      36.4%; Score 364; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.5e-84;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTCACCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCC 60
Db 364 ACACAGTCACCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCC 305

QY 61 CAAGACCCCTTCCTGTCCTCCCTACCTCATCTGACCTTTCCACCTCCCTCCAGGTCCAATGCC 120
Db 304 CAAGACCCCTTCCTGTCCTCCCTACCTCATCTGACCTTTCCACCTCCCTCCAGGTCCAATGCC 245

QY 121 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCCGTATGAACCAATTCAGCCCC 180
```

```
Db 244 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCCGTATGAACCAATTCAGCCCC 185

QY 181 TACTGTCCCGTGTTCGGCAATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCAGGAC 240
Db 184 TACTGTCCCGTGTTCGGCAATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCAGGAC 125

QY 241 CTGGCGTTCTGCTGGTCCCAAGTTGGGGCAGAGTTCCTTAGAGGGTCTTGGGAGAGGG 300
Db 124 CTGGCGTTCTGCTGGTCCCAAGTTGGGGCAGAGTTCCTTAGAGGGTCTTGGGAGAGGG 65

QY 301 TCCGGGGCCACCCACCGGTGAAAAGCTATGTCTATGTGACAGGTGGTCTGTAGGCA 360
Db 64 TCCGGGGCCACCCACCGGTGAAAAGCTATGTCTATGTGACAGGTGGTCTGTAGGCA 5

QY 361 TCAG 364
Db 4 TCAG 1

RESULT 13
CQ093442/c
LOCUS      440 bp      DNA      linear      PAT 21-JAN-2004
DEFINITION Sequence 2301 from Patent WO0157272.
ACCESSION CQ093442
VERSION    CQ093442.1 GI:41062468
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
            Patent: WO 0157272-A 2301 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES   Location/Qualifiers
            source
            1..440
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="MAP TO AC002472.3-EXPRESSED IN PLACENTA, SIGNAL = 6"

ORIGIN
Query Match      36.4%; Score 364; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.5e-84;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACACAGTCACCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCC 60
Db 364 ACACAGTCACCTTCAGCAAGTTCAACTTCTTAAGTAAGCAGAGTGGTCTCATCTGCC 305

QY 61 CAAGACCCCTTCCTGTCCTCCCTACCTCATCTGACCTTTCCCACTCCCTCCAGGTCCAATGCC 120
Db 304 CAAGACCCCTTCCTGTCCTCCCTACCTCATCTGACCTTTCCCACTCCCTCCAGGTCCAATGCC 245

QY 121 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCCGTATGAACCAATTCAGCCCC 180
Db 244 TTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCCGTATGAACCAATTCAGCCCC 185

QY 181 TACTGTCCCGTGTTCGGCAATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCAGGAC 240
Db 184 TACTGTCCCGTGTTCGGCAATTTGGGACCTCTGTCGCAAGCTTGGAGGACCTTCAGGAC 125

QY 241 CTGGCGTTCTGCTGGTCCCAAGTTGGGGCAGAGTTCCTTAGAGGGTCTTGGGAGAGGG 300
Db 124 CTGGCGTTCTGCTGGTCCCAAGTTGGGGCAGAGTTCCTTAGAGGGTCTTGGGAGAGGG 65

QY 301 TCCGGGGCCACCCACCGGTGAAAAGCTATGTCTATGTGACAGGTGGTCTGTAGGCA 360
Db 64 TCCGGGGCCACCCACCGGTGAAAAGCTATGTCTATGTGACAGGTGGTCTGTAGGCA 5
```



```
QY 361 TCAG 364
Db 4 TCAG 1

RESULT 14
LOCUS CQ132213/c
DEFINITION Sequence 2235 from Patent WO0157276.
ACCESSION CQ132213
VERSION CQ132213.1 GI:41089569
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human bone marrow
Patent: WO 0157276-A 2235 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source
1..440
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002472.3-EXPRESSED IN BONE MARROW, SIGNAL
= 3.6"

ORIGIN
Query Match 36.4%; Score 364; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.5e-84;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACACAGTCACCTTCAGCAAGTTCACCTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC 60
Db 364 ACACAGTCACCTTCAGCAAGTTCACCTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC 305
QY 61 CAAGACCTCTCTGTCCTCCCTACCTCATCTGACCTTCCACCTCTCCAGGTCCCAATGCC 120
Db 304 CAAGACCTCTCTGTCCTCCCTACCTCATCTGACCTTCCACCTCTCCAGGTCCCAATGCC 245
QY 121 TTGGAGACTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180
Db 244 TTGGAGACTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 185
QY 181 TACTGTCCCGTGTTCGGCATTTGGGACCTTCGTGGCCAAAGCTTGGAGGACCTTCGAGGAC 240
Db 184 TACTGTCCCGTGTTCGGCATTTGGGACCTTCGTGGCCAAAGCTTGGAGGACCTTCGAGGAC 125
QY 241 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGCTCTGGAGAGGG 300
Db 124 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGCTCTGGAGAGGG 65
QY 301 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGCAGGCTGGCTCTGTAGGCA 360
Db 64 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGCAGGCTGGCTCTGTAGGCA 5
QY 361 TCAG 364
Db 4 TCAG 1

Search completed: November 22, 2004, 01:15:08
Job time : 4288.34 secs

QY 361 TCAG 364
Db 4 TCAG 1

RESULT 15
LOCUS CQ170783/c
DEFINITION Sequence 2179 from Patent WO0157274.
ACCESSION CQ170783
VERSION CQ170783.1 GI:41165519
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human bone marrow
Patent: WO 0157276-A 2235 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source
1..440
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002472.3-EXPRESSED IN BONE MARROW, SIGNAL
= 3.6"

ORIGIN
Query Match 36.4%; Score 364; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.5e-84;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACACAGTCACCTTCAGCAAGTTCACCTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC 60
Db 364 ACACAGTCACCTTCAGCAAGTTCACCTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC 305
QY 61 CAAGACCTCTCTGTCCTCCCTACCTCATCTGACCTTCCACCTCTCCAGGTCCCAATGCC 120
Db 304 CAAGACCTCTCTGTCCTCCCTACCTCATCTGACCTTCCACCTCTCCAGGTCCCAATGCC 245
QY 121 TTGGAGACTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180
Db 244 TTGGAGACTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 185
QY 181 TACTGTCCCGTGTTCGGCATTTGGGACCTTCGTGGCCAAAGCTTGGAGGACCTTCGAGGAC 240
Db 184 TACTGTCCCGTGTTCGGCATTTGGGACCTTCGTGGCCAAAGCTTGGAGGACCTTCGAGGAC 125
QY 241 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGCTCTGGAGAGGG 300
Db 124 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGCTCTGGAGAGGG 65
QY 301 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGCAGGCTGGCTCTGTAGGCA 360
Db 64 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGCAGGCTGGCTCTGTAGGCA 5
QY 361 TCAG 364
Db 4 TCAG 1

RESULT 15
LOCUS CQ170783/c
DEFINITION Sequence 2179 from Patent WO0157274.
ACCESSION CQ170783
VERSION CQ170783.1 GI:41165519
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human heart
Patent: WO 0157274-A 2179 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
source
1..440
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC002472.3-EXPRESSED IN HEART, SIGNAL = 6.2"

ORIGIN
Query Match 36.4%; Score 364; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.5e-84;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACACAGTCACCTTCAGCAAGTTCACCTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC 60
Db 364 ACACAGTCACCTTCAGCAAGTTCACCTTCTCTAAGTAAGCAGAGTGGGTCTCATCTGCCC 305
QY 61 CAAGACCTCTCTGTCCTCCCTACCTCATCTGACCTTCCACCTCTCCAGGTCCCAATGCC 120
Db 304 CAAGACCTCTCTGTCCTCCCTACCTCATCTGACCTTCCACCTCTCCAGGTCCCAATGCC 245
QY 121 TTGGAGACTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 180
Db 244 TTGGAGACTGGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAGCCCC 185
QY 181 TACTGTCCCGTGTTCGGCATTTGGGACCTTCGTGGCCAAAGCTTGGAGGACCTTCGAGGAC 240
Db 184 TACTGTCCCGTGTTCGGCATTTGGGACCTTCGTGGCCAAAGCTTGGAGGACCTTCGAGGAC 125
QY 241 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGCTCTGGAGAGGG 300
Db 124 CTGGCGTTGCTGGTGGTCCCAAGTTGGGGCAGGGTTCCTAGAGGCTCTGGAGAGGG 65
QY 301 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGCAGGCTGGCTCTGTAGGCA 360
Db 64 TCCCGGGCCACCCACCGGTGGAAGCTATGTCTATGTGCAGGCTGGCTCTGTAGGCA 5
QY 361 TCAG 364
Db 4 TCAG 1

Search completed: November 22, 2004, 01:15:08
Job time : 4288.34 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: November 21, 2004, 14:42:41 ; Search time 191 Seconds  
(without alignments)  
1220.035 Million cell updates/sec

Title: US-09-820-095B-2  
Perfect score: 2226  
Sequence: 1 MGSPGATTGWLIDYKTEKW.....TPGWPCPSDRHLPTHGSSL 405

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2203	99.0	431	1 P2X6 HUMAN	O15547 homo sapien
2	2203	99.0	441	2 Q61C33	Q61c33 homo sapien
3	2203	99.0	441	2 CAG30421	Cag30421 homo sapi
4	1599	71.8	379	1 P2X6 MOUSE	O54803 mus musculu
5	1554	69.8	379	1 P2X6 RAT	P51579 rattus norv
6	946	42.5	402	2 Q90W80	Q90w80 gallus gall
7	943	42.4	402	2 Q91AD2	Q91ad2 gallus gall
8	900.5	40.5	409	2 Q90X65	Q90x65 rana catesb
9	892	40.1	455	1 P2X5 RAT	P51578 rattus norv
10	888	39.9	455	2 Q91VE2	Q91ve2 mus musculu
11	885.5	39.8	481	2 Q8JFP7	Q8jfp7 brachydanio
12	822	36.9	422	2 Q81XW4	Q81xw4 homo sapien
13	816	36.7	388	1 P2X4 HUMAN	Q99571 homo sapien
14	813	36.5	388	2 Q8NA1	Q8n4n1 homo sapien
15	803.5	36.1	421	1 P2X5 HUMAN	Q93086 homo sapien
16	795	35.7	388	2 Q92257	Q92257 mus musculu
17	795	35.7	388	2 Q9JUX6	Q9jux6 mus musculu
18	794.5	35.7	385	2 Q9YI70	Q9yi70 gallus gall
19	793	35.6	388	1 P2X4 RAT	P51577 rattus norv
20	789	35.4	384	2 Q9PU37	Q9pu37 gallus gall
21	788	35.4	391	2 Q9DDP0	Q9ddp0 xenopus lae
22	785	35.3	384	2 Q9DDP3	Q9ddp3 xenopus lae
23	785	35.3	391	2 Q9DDP1	Q9ddp1 xenopus lae
24	777	34.9	391	2 Q9DDP1	Q9ddp1 xenopus lae
25	770	34.6	391	2 Q9DDP2	Q9ddp2 xenopus lae
26	755	33.9	389	2 Q6NVR1	Q6nvr1 brachydanio
27	755	33.9	389	2 AAH6495	AAh6495 brachydanio
28	754	33.9	389	2 Q9ETZ0	Q9etz0 brachydanio
29	744	33.4	399	1 P2X1 HUMAN	P51575 mus sapien
30	740	33.2	399	1 P2X1 MOUSE	P51576 mus musculu
31	740	33.2	399	2 Q91W13	Q91w13 mus musculu

ALIGNMENTS

RESULT 1

ID	P2X6 HUMAN	STANDARD;	PRT;	431 AA.
AC	O15547; Q9UL50;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	P2X purinoceptor 6 (ATP receptor) (P2X6) (Purinoergic receptor) (P2XM)			
DE	(Purinoergic receptor P2X-like 1).			
GN	Name=P2RX11; Synonyms=P2RX6, P2X6;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97384966; PubMed=9242461;			
RA	Urano T., Nashimori H., Han H., Furuhata T., Kimura Y., Nakamura Y.,			
RA	Tokino T.;			
RT	"Cloning of P2XM, a novel human P2X receptor gene regulated by p53.;"			
RL	Cancer Res. 57:3281-3287(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skeletal muscle;			
RA	Cheng X., Jin H., Huang C.-C.;			
RT	"Cloning and tissue distribution of a human cDNA encoding P2X6			
RL	purinoceptor.;"			
RL	Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	PHOSPHORYLATION SITE TYR-54.			
RX	MEDLINE=22107313; PubMed=12112843;			
RA	DOI=10.1002/1519-9861(200206)2:5<642::AID-PROT642>3.0.CO;2-I;			
RA	Maguire P.B., Wynne K.J., Harney D.F., O'Donoghue N.M., Stephens G.,			
RA	Fitzgerald D.J.;			
RT	"Identification of the phosphotyrosine proteome from thrombin			
RL	activated platelets.;"			
RL	Proteomics 2:642-648(2002).			
CC	-1- FUNCTION: Receptor for ATP that acts as a ligand gated ion channel			
CC	(By similarity).			
CC	-1- SUBUNIT: Homo- or heteropolymers (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- TISSUE SPECIFICITY: Expressed predominantly in skeletal muscle.			
CC	-1- SIMILARITY: Belongs to the P2X receptor family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; AB002059; BAA22047.1; -			
DR	EMBL; AB002058; BAA22046.1; -			

Aaf68968 mus muscu  
Bac29024 mus muscu  
F47824 rattus norv  
Aah61742 rattus no  
O70397 cavia porce  
Q8awd8 brachydanio  
Q8awc0 gallus gall  
Q7szl7 brachydanio  
Q78dw2 rattus norv  
Q7szm0 brachydanio  
F49653 rattus norv  
O70399 cavia porce  
O8k3p1 mus musculu  
Q78dw3 rattus norv

DR	ENBL; AF065385; AAF13303.1; -.
DR	Genew; HGNC:8538; P2RXL1.
DR	MIM; 608077; -.
DR	GO; GO:0005887; C:integral to plasma membrane; TAS.
DR	GO; GO:0015267; F:channel/pore class transporter activity; TAS.
DR	GO; GO:0004888; F:transmembrane receptor activity; TAS.
DR	GO; GO:006936; P:muscle contraction; TAS.
DR	GO; GO:0007165; P:signal transduction; TAS.
DR	GO; GO:0006810; P:transport; TAS.
DR	Pfam; PF00864; P2X_receptor.1.
DR	PRINTS; PR01307; P2XRECEPTOR.
DR	TIGRFAMS; TIGR00863; P2X; 1.
DR	PROSITE; PS01212; P2X RECEPTOR; 1.
KW	Glycoprotein; Ion transport; Ionic channel; Phosphorylation; Receptor;
KW	transmembrane.
FT	DOMAIN 1 29 Cytoplasmic (Potential).
FT	TRANSMEM 30 50 1 (Potential).
FT	DOMAIN 51 323 Extracellular, cysteine-rich (Potential).
FT	TRANSMEM 324 344 2 (Potential).
FT	DOMAIN 345 431 Cytoplasmic (Potential).
FT	MCD_RES 54 54 Phosphotyrosine.
FT	CARBOHYD 155 155 N-linked (GlcNAc.. ) (Potential)..
FT	CARBOHYD 185 185 N-linked (GlcNAc.. ) (Potential)..
FT	CARBOHYD 200 200 N-linked (GlcNAc.. ) (Potential).
FT	CONFLICT 383 383 L -> F (in Ref. 2).
SQ	SEQUENCE 431 AA; 47912 MW; ODE1920178225F3B CRC64;
Query Match 99.0%; Score 2203; DB 1; Length 431;	
Best Local Similarity 94.0%; Pred.No. 2.8e-193;	
Matches 405; Conservative 0; Mismatches 0; Indels 26; Gaps 1;	
QY	1 MGSPGATTGWGLLDYKTEK-----WALLAKKGQERDL 34
Db	1 MGSPGATTGWGLLDYKTEKYVMTNRWRVGALQRLLQGIVVVGVGWallakkgqerdl 60
QY	35 PFSIITKLKVSVTQIKELGNRLWDADVFKVPQGENVFVLVTFLTPAQVQCPEH 94
Db	61 PFSIITKLKVSVTQIKELGNRLWDADVFKVPQGENVFVLVTFLTPAQVQCPEH 120
QY	95 PSVPLANCWDEDCEGEGETSHSGVTKGCVFNFGTHRTCTISWCWPVESVPSRLL 154
Db	121 PSVPLANCWDEDCEGEGETSHSGVTKGCVFNFGTHRTCTISWCWPVESVPSRLL 180
QY	155 AQAOFTLFINKTVTFSKNFPSKSNALETWDPITYFKHCERYEQFPYCVPFRIGDLVAKA 214
Db	181 AQAOFTLFINKTVTFSKNFPSKSNALETWDPITYFKHCERYEQFPYCVPFRIGDLVAKA 240
QY	215 GTTFEDLALGGSGVICRVHDCDLLDTGDSCWPHYSFQLQESYNFRATATHWEOPVEA 274
Db	241 GTTFEDLALGGSGVICRVHDCDLLDTGDSCWPHYSFQLQESYNFRATATHWEOPVEA 300
QY	275 RTLLKLYGIREDILVTQAOKFGLIPFAVTLGTGAAWLGWTFFCDDLILLYYVDREAHPYW 334
Db	301 RTLLKLYGIREDILVTQAOKFGLIPFAVTLGTGAAWLGWTFFCDDLILLYYVDREAHPYW 360
QY	335 RTKYEEAKAPATANSVMRELALASQARLAECILRRSSAPAPTATAAGSQTOTPGMPCPSS 394
Db	361 RTKYEEAKAPATANSVMRELALASQARLAECILRRSSAPAPTATAAGSQTOTPGMPCPSS 420
QY	395 DTHLPETHSGSL 405
Db	421 DTHLPETHSGSL 431

## RESULT 2

RESIDUE 2	PRELIMINARY;	PRT;	441 AA.
Q6IC33			
AD			
IC Q6IC33;			
DT 05-JUL-2004	(TREMblrel. 27,		Created)
DT 05-JUL-2004	(TREMblrel. 27,		Last sequence update)
DT 05-JUL-2004	(TREMblrel. 27,		Last annotation update)
DE P2RXL1	protein.		

GN	Name=P2RXL1;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC	NCBI_TaxID=9606;
[1]	
RN	SEQUENCE FROM N.A.
RP	Collins J.B., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,
RA	Cole C.G., Goward M.E., Aguado B., Malliya M., Motrab Y., Huckle E.J.,
RA	Beare D.M., Dunham I.;
RL	Submitted (May-2004) to the EMBL/GenBank/DBJ databases.
CC	- - SUBUNIT: Homo- or heteropolymers (By similarity).
CC	- - SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC	- - SIMILARITY: Belongs to the P2X receptor family.
DR	EMBL; CR456533; CAG30421.1; -;
DR	InterPro; IPR003049; P2X6_purinocptor.
DR	InterPro; IPR001429; P2X_receptor.
DR	Pfam; PR00864; P2X_receptor.1.
DR	PRINTS; PR01313; P2X6RECEPTOR.
DR	PRINTS; PR01307; P2XRECEPTOR.
DR	TIGRFAMS; TIGR00863; P2X.1.
DR	FROSTITE; PS01212; P2X_RECEPTOR.1.
KW	Glycoprotein; Ion transport; Ionic channel; Receptor; Transmembrane;
KW	Transport.
SQ	SEQUENCE 441 AA; 48928 MW; EFD45BDC81234A64 CRC64;
	Query Match 99.0%; Score 2203; DB 2; Length 441;
	Best Local Similarity 94.0%; Pred. No. 2.9e-193;
	Matches 405; Conservative 0; Mismatches 0; Indels 26; Gaps 1
Qy	1 MGSPGATTGWGLDYKTEK-----WALLAKKGYQERDLE 34
Db	11 MGSPGATTGWGLLDYKTEKYVMTENRWVGALQRLLQFGIVVYVGVWALLAKKGYQERDLE 70
Qy	35 POFSIITKLGVSVTQIKELGNRLNDVADFVKPSQGENVFVLVTNLVTPAQVQCRCPEH 94
Db	71 POFSIITKLGVSVTQIKELGNRLNDVADFVKPSQGENVFVLVTNLVTPAQVQCRCPEH 130
Qy	95 PSVPPLANCWDEDCEPEGEGTHSGVKTGCVVFNFTHRTRCEIWSVCPVESGWVPSRPLL 154
Db	131 PSVPPLANCWDEDCEPEGEGTHSGVKTGCVVFNFTHRTRCEIWSVCPVESGWVPSRPLL 190
Qy	155 AQAQNFTLIIFIKNTVTFSKFNFSKNALETWDPTFYFKCHRYEQFSPCYCVPIRGDLVAKA 214
Db	191 AQAQNFTLIIFIKNTVTFSKFNFSKNALETWDPTFYFKCHRYEQFSPCYCVPIRGDLVAKA 250
Qy	215 GTGFEDLALLGGSGVIRVHWDCDLLDTGDSGCWPHYSFQLQEKSYNPRTATHWWEPQVEA 274
Db	251 GTGFEDLALLGGSGVIRVHWDCDLLDTGDSGCWPHYSFQLQEKSYNPRTATHWWEPQVEA 310
Qy	275 RTLLKLYGIRDIILVTQAOKFGLIPPAVTLTGTGAWLGVWTFFCDLLLLYYVDRAHFYW 334
Db	311 RTLLKLYGIRDIILVTQAOKFGLIPPAVTLTGTGAWLGVWTFFCDLLLLYYVDRAHFYW 370
Qy	335 RTKYEEAKAPATANSVWRRELASQARLAECRLRSAPAPTATAAGSQTOTPGWPCCSS 394
Db	371 RTKYEEAKAPATANSVWRRELASQARLAECRLRSAPAPTATAAGSQTOTPGWPCCSS 430
Qy	395 DTHLPETHSGSL 405
Db	431 DTHLPETHSGSL 441

### RESULT 3

CA3030421	PRELIMINARY;	PRT;	441 AA.
ID	CAG30421		
AC	CAG30421;		
DT	01-JUN-2004 (TrEMBLrel. 27, Created)		
DT	01-JUN-2004 (TrEMBLrel. 27, Last sequence update)		
DT	01-JUN-2004 (TrEMBLrel. 27, Last annotation update)		
DE	P2RXL1 protein.		
DE	P2RXL1.		
GN	P2RXL1.		
OS	Homo sapiens (Human).		

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Collins J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,  
 RA Cole C.G., Ward M.E., Aguado B., Mallia M., Mokrab Y., Huckle E.J.,  
 RA Beare D.M., Dunham I.,  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR456535; CAG30421.1; ..  
 SQ SEQUENCE 441 AA; 48828 MW; EFD45BDC81234A64 CRC64;

Query Match 99.0%; Score 2203; DB 2; Length 441;  
 Best Local Similarity 94.0%; Pred. No. 2.9e-193;  
 Matches 405; Conservative 0; Mismatches 0; Indels 26; Gaps 1;

QY 1 MGSPGATTGGLLDYKTEK-----WALLAKGYQERDLE 34  
 Db 11 MGSPGATTGGLLDYKTEKVMRWVGALQELQFGIVVVGWALLAKGYQERDLE 70  
 QY 35 PQSIIITKLKGVSVTOIKELGNRLWDVADFPKPPQGENVFFLVNPLVTPAQVQGRCPBH 94  
 Db 71 PQSIIITKLKGVSVTOIKELGNRLWDVADFPKPPQGENVFFLVNPLVTPAQVQGRCPBH 130  
 QY 95 PSVPLANCWVDEDCPEGEGETSHGVKTCQCVFNGTHTCTCIWSKCPVSGVPSRPL 154  
 Db 131 PSVPLANCWVDEDCPEGEGETSHGVKTCQCVFNGTHTCTCIWSKCPVSGVPSRPL 190  
 QY 155 AQQNFTLFIKNTVTFESKFNFSKNALETWDTYFKHCYERQFPSPYCPVFRIGDLVAKA 214  
 Db 191 AQQNFTLFIKNTVTFESKFNFSKNALETWDTYFKHCYERQFPSPYCPVFRIGDLVAKA 250  
 QY 215 GGTFFDLALGSGVGRVHWDCLDTGDSGCPHYSFQLEKSYNFRATHWQPGVBA 274  
 Db 251 GGTFFDLALGSGVGRVHWDCLDTGDSGCPHYSFQLEKSYNFRATHWQPGVBA 310  
 QY 275 RTLLKLYGRFDILVTGQAGKFLIPTAVTLGTGAALGWTFPFCDDLLLYVDREAHFW 334  
 Db 311 RTLLKLYGRFDILVTGQAGKFLIPTAVTLGTGAALGWTFPFCDDLLLYVDREAHFW 370  
 QY 335 RTKYBAKAPKATANSVWRELALASQARLAECRLRSSAPATATAAGSQTQTPGWPSPSS 394  
 Db 371 RTKYBAKAPKATANSVWRELALASQARLAECRLRSSAPATATAAGSQTQTPGWPSPSS 430  
 QY 395 DTHLPHTHSGSL 405  
 Db 431 DTHLPHTHSGSL 441

## RESULT 4

P2X6 MOUSE  
 ID P2X6 MOUSE STANDARD; PRT; 379 AA.  
 AC O54803;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE P2X purinoceptor 6 (ATP receptor) (P2X6) (Purinergic receptor) (P2XM)  
 DE (Purinergic receptor P2X-like 1).  
 GN Name=P2rx11; Synonyms=P2rx6;  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069782; PubMed=9852680;  
 RA Nawa G., Urano T., Tokino T., Ochi T., Miyoshi Y.;  
 RT "Cloning and characterization of the murine P2XM receptor gene";  
 RL J. Hum. Genet. 43:262-267(1998).  
 CC -!- FUNCTION: Receptor for ATP that acts as a ligand gated ion channel  
 CC (By similarity).  
 CC -!- SUBUNIT: Homo- or heteropolymers (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- TISSUE SPECIFICITY: Predominantly expressed in skeletal muscle.  
 CC Also expressed in lung.  
 CC -!- SIMILARITY: Belongs to the P2X receptor family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB010883; BAA24693.1; ..  
 DR MGI; 1337113; P2rx11.  
 DR InterPro; IPR001429; P2X receptor.  
 DR Pfam; PF00864; P2X receptor; 1.  
 DR PRINTS; PR01307; P2XRECEPTOR.  
 DR TIGRfams; TIGR00863; P2X; 1.  
 DR PROSITE; PS01212; P2X\_RECEPTOR; 1.  
 KW Glycoprotein; Ion transport; Ionic channel; Receptor; Transmembrane.  
 FT DOMAIN 1 31  
 FT TRANSMEM 32 52  
 FT DOMAIN 53 325  
 FT TRANSMEM 326 346  
 FT DOMAIN 347 379  
 FT CARBOHYD 157 187  
 FT CARBOHYD 187 187  
 FT CARBOHYD 202 202  
 SQ SEQUENCE 379 AA; 42115 MW; 1P2CF565934133E5 CRC64;

Query Match 71.8%; Score 1599; DB 1; Length 379;  
 Best Local Similarity 78.7%; Pred. No. 6.4e-138;  
 Matches 292; Conservative 17; Mismatches 36; Indels 26; Gaps 1;

QY 6 ATTGGLLDYKTEK-----WALLAKGYQERDLEPQFSI 39  
 Db 8 ALVSMGFLDYKTEKXVLTNCRVGVSRLLQLAVVVVIGWALLAKGYQERDLAPQTSV 67  
 QY 40 ITKLKGVSVTOIKELGNRLWDVADFPKPPQGENVFFLVNPLVTPAQVQGRCPBHPSVPL 99  
 Db 68 ITKLKGVSVTOIKELGNRLWDVADFPKPPQGENVFFLVNPLVTPAQVQGRCPBHPSVPL 127  
 QY 100 ANCWVDSDCEPEGEGETSHGVKTCQCVFNGTHTCTCIWSKCPVSGVPSRPLAQON 159  
 Db 128 ANCWVDSDCEPEGEGETSHGVKTCQCVFNGTHTCTCIWSKCPVSGVPSRPLAQON 187  
 QY 160 PTLFIKNTVTFESKFNFSKNALETWDTYFKHCYERQFPSPYCPVFRIGDLVAKAGTPE 219  
 Db 188 PTLFIKNTVTFESKFNFSRNNALLTWDNTYFKHCYERQFPSPYCPVFRIGDLVAKAGTPE 247  
 QY 220 DLALLGSGVGRVHWDCLDTGDSGCPHYSFQLEKSYNFRATHWQPGVBAEARTLLK 279  
 Db 248 DLALLGSGVGRVHWDCLDTGDSGCPHYSFQLEKSYNFRATHWQPGVBAEARTLLK 307  
 QY 280 LYGRFDILVTGQAGKFLIPTAVTLGTGAALGWTFPFCDDLLLYVDREAHFWRTKYE 339  
 Db 308 LYGRFDILVTGQAGKFLIPTAVTLGTGAALGWTFPFCDDLLLYVDREAHFWRTKYE 367  
 QY 340 EAKAPKATANS 350  
 Db 368 EAKAPKATANS 378

RESULT 5  
 P2X6 RAT  
 ID P2X6 RAT STANDARD; PRT; 379 AA.  
 AC P51579;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE P2X purinoceptor 6 (ATP receptor) (P2X6) (Purinergic receptor) (P2XM)  
 DE (Purinergic receptor P2X-like 1).  
 GN Name=P2rx11; Synonyms=P2rx6;





QY 165 KNTVTSKFNFSKNALETWPTYPFKHCRYPQSPFCVFRIGDLVAKAGTGFEDLALL 224  
 DB 193 KNSIRFPKFNFSKNTVRDTKGSPLKTCRHSKD-DHYCPIPHLGKIVSWSGDFODIALE 251  
 QY 225 GSGVIRVHWDCLDTGSGCWPHYSF-----QIQEKS-----YNFRATHWEOFGVEAR 275  
 DB 252 GGVIGIQIEWDCNLDKDASECHPRYSFTRLDNKFSEKSVSGYNFRYAKYRDVNETDYR 311  
 QY 276 TLLKLYGRFDFILVTGQAGKFLIPTAVTLGTGAALWGVTFPFDLLLLYVDREAHFVWR 335  
 DB 312 TLIKAYGIRFDIWMVGKAGKFNIIITINISGLALMGAGAFFCDLVLLYVYKRSFYRD 371  
 QY 336 KYEBEAKAPK 345  
 DB 372 KKEFEVKSVK 381

RESULT 9  
 P2X5\_RAT  
 ID P2X5\_RAT STANDARD; PRT; 455 AA.  
 AC P51578; Q64613;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE P2X5 purinoceptor 5 (ATP receptor) (P2X5) (Purinergic receptor).  
 GN Name=P2rx5;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Coeliac ganglion;  
 RX MEDLINE=96256686; PubMed=8786426;  
 RA Collo G.H., Kawashima B., Pich E., Neidhart S., North R.A.,  
 RA Surprenant A., Buell G.N.;  
 RT "Cloning of P2X5 and P2X6 receptors and the distribution and  
 RT properties of an extended family of ATP-gated ion channels";  
 RL J. Neurosci. 16:2495-2507(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Heart;  
 RX MEDLINE=96275514; PubMed=8690069;  
 RA Garcia-Guzman M., Soto F., Laube B., Stuehmer W.;  
 RT "Molecular cloning and functional expression of a novel rat heart P2X  
 RT purinoceptor";  
 RL FEBS Lett. 388:123-127(1996).  
 CC -!- FUNCTION: Receptor for ATP that acts as a ligand gated ion  
 CC channel.  
 CC -!- SUBUNIT: Homo- or heteropolymers (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Predominantly expressed in heart but are also  
 CC present in brain, spinal cord and adrenal gland.  
 CC -!- SIMILARITY: Belongs to the P2X receptor family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC  
 CC EMBL; X92069; CAA63052.1; -  
 CC EMBL; X93328; CAA65993.1; -  
 CC PIR; S71344; S71344.  
 CC RGD; 620256; P2rx5.  
 CC InterPro; IPR001429; P2X\_receptor.  
 CC Pfam; PF00864; P2X\_receptor; 1.  
 CC PRINTS; PR01307; P2XRECEPTOR.  
 CC TIGRFAMs; TIGR00863; P2X; 1.  
 CC PROSITE; P501212; P2X\_RECEPTOR; 1.

KW Glycoprotein; Ion transport; Ionic channel; Receptor; Transmembrane.  
 FT DOMAIN 1 30 Cytoplasmic (Potential).  
 FT TRANSMEM 31 51 1 (Potential).  
 FT DOMAIN 52 341 Extracellular, cysteine-rich (Potential).  
 FT TRANSMEM 342 362 2 (Potential).  
 FT DOMAIN 363 455 Cytoplasmic (Potential).  
 FT CARBOHYD 77 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 157 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 202 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 191 S -> F (in Ref. 2).  
 FT CONFLICT 396 R -> Q (in Ref. 2).  
 FT SEQUENCE 455 AA; 51479 MW; D4405BB440490F70 CRC64;  
 Query Match 40.1%; Score 892; DB 1; Length 455;  
 Best Local Similarity 46.6%; Pred. No. 66-73;  
 Matches 174; Conservative 55; Mismatches 108; Indels 36; Gaps 5;  
 QY 12 LLDYKTEK-----WALLAKGYQBRDLEPQPSIITKLKG 45  
 DB 13 LPDYKTAKFVVAKSKVGLLYRVLQLILLYLLIWWFLIKSYQDIDTSLQSAVTVKVG 72  
 QY 46 VSVTQIKELGNRLWADVVPKPPQGENVFLLVTNFLTPTPAQVQGRCPHPSPVLANCWD 105  
 DB 73 VAYTNTMLGERLWADVFPISQGENVFVVTNLTVPNORQGI CAEREGIPDGECSBD 132  
 QY 106 EDCPEGEGETHSHGVKTCGV-VFNGTHTCTCIWMCVPVSGVSPRLPAAQNFILFI 164  
 DB 133 DCHAGESVVAAGLTKTGRVGNSTRCTGCLFANCPVETKSMPTDPLLKDAESFTISI 192  
 QY 165 KNTVTSKFNFSKNALETWPTYPFKHCRYPQSPFCVFRIGDLVAKAGTGFEDLALL 224  
 DB 193 KXFIKPFKFNFSKANVLETDNKHFLKTCFHS-STNLYCPIFRLGSIVRWAGADFQIALK 251  
 QY 225 GGSVGRVHWDCLDTGDSGCWPHYSF-QIQEK-----SYNFRATHWEOFGVEAR 276  
 DB 252 GGVIGIQIEWDCNLDKDAKSCNPNFYFNLDNKNHTSSSGNFRFPARYRDPNGVEFRD 311  
 QY 277 LKLYGIRFDILVTGQAGKFLIPTAVTLGTGAALWGVTFPFDLLLLYVDREAHFVWR 336  
 DB 312 LMKAYGIRFDIVNGKAGKFSIPTVINIGSLALMGAGAFFCDLVLLYVYKRSFYRDK 371  
 QY 337 KYEBEAKAPKATAN 349  
 DB 372 KKEKVGQKEDAN 384

RESULT 10  
 Q91VE2  
 ID Q91VE2 PRELIMINARY; PRT; 455 AA.  
 AC Q91VE2;  
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
 DE ATP-gated ionotropic P2X5 receptor subunit.  
 GN Name=P2rx5;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvJ, and BALB/c; TISSUE=Heart;  
 RX MEDLINE=21297188; PubMed=11404011;  
 RA Cox J.A., Barmina O., Voigt M.M.;  
 RT "Gene structure, chromosomal localization, cDNA cloning and expression  
 RT of the mouse ATP-gated ionotropic receptor P2X5 subunit";  
 RL Gene 270:145-152(2001).  
 CC -!- SUBUNIT: Homo- or heteropolymers (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to the P2X receptor family.  
 CC EMBL; AF333333; AAK49937.1; -  
 CC EMBL; AF333332; AAK49937.1; JOINED.  
 CC EMBL; AF333331; AAK49936.1; -.

DR	MGI:2137026; P2rx5.	DR	EMBL; AF500300; AAM74150.1; --
DR	GO: 0005887; C: integral to plasma membrane; IC.	DR	ZFIN; ZDB-GENE-030319-5; P2rx5
DR	GO: 0004931; F: ATP-gated cation channel activity; IDA.	DR	GO: 00016021; C: integral to membrane; IEA.
DR	InterPro: IPR001429; P2X5_purinocceptor.	DR	GO: 0005524; F: ATP binding; IEA.
DR	Pfam: PF00864; P2X receptor.	DR	GO: 0005216; F: ion channel activity; IEA.
DR	PRINTS; PR01307; P2X5RECEPTOR.	DR	GO: 0004872; F: receptor activity; IEA.
DR	PRINTS; PR01307; P2X5RECEPTOR.	DR	GO: 0006811; F: ion transport; IEA.
DR	TIGRFAMs; TIGR00863; P2X; 1.	DR	InterPro: IPR001429; P2X receptor.
DR	TIGRFAMs; TIGR00863; P2X; 1.	DR	Pfam: PF00864; P2X receptor; 1.
DR	PROSITE; PS01212; P2X RECEPTOR; 1.	DR	PRINTS; PR01307; P2X5RECEPTOR.
KW	Glycoprotein; Ion transport; Ionic channel; Receptor; Transmembrane;	DR	TIGRFAMs; TIGR00863; P2X; 1.
KW	Transport.	DR	PROSITE; PS01212; P2X RECEPTOR; 1.
SQ	SEQUENCE 455 AA; 51461 MW; 0DC09DEB131CAEB0 CRC64;	KW	Glycoprotein; Ion transport; Ionic channel; Receptor; Transmembrane;
Query Match 39.8%; Score 888; DB 2; Length 455;			
Best Local Similarity 42.2%; Pred. No. 1.4e-72;			
Matches 175; Conservative 54; Mismatches 114; Indels 42; Gaps 5;			
QY	12 LLDYKTEK-----	QY	10 WGLDYKTEK-----
DB	13 LFDYKAKFVAKSKVGLLYRVLQTLILYLLIWLFLIKSYQDIDTSLQSAVTVKVG 72	DB	10 FSLDYKTEKFIKAKKVGVLFRFLQTLTGIGYLGWVFIWKGYQETEEAIQSSVITKL 69
QY	46 VSVTQIKELGNRLWDVADFKPPQGENVFPLVTFNLTVAQVQGRCPPEHSPVLANCWD 105	QY	44 KGVSVTQIKELGNRLWDVADFKPPQGENVFPLVTFNLTVAQVQGRCPPEHSPVLANCW 103
DB	73 VAYNTTTLGERLWDVADFKVPSQGENVFVVTNLTVPNQRGICAEREGIPDGCESD 132	DB	70 KGVDLTNSQFGLQWGAEDYVIPPQGRDVFVVTNLTVPNQRHGHCPESPVPDGFCT 129
QY	106 EDCPEGGTHSHGVITGQCV-VFNGTHRTCEIWSVCPVESGVVPSRPLLAQAQNPLFI 164	QY	104 VDEDCPEGGTHSHGVITGQCVVFNGTHRTCEIWSVCPVESGVVPSRPLLAQAQNPLFI 163
DB	133 TDCHAGESVAGHGLATGRCLRYGNSGTGTCEIFAMCPVETKSMPTDPLDKDAEGFTIF 192	DB	130 NDNECVGESVLAGLVKVTGRCLNDTG---TCEINAWCPVEHGHAPVEPLAKAENFTVI 186
QY	165 KNTVTSKFNFSKNALETWDPYFKHCRYEPQSPYCPVERIGDLVAKAGGTFFEDLALL 224	QY	164 KNTVTSKFNFSKNALETWDPYFKHCRYEPQSPYCPVERIGDLVAKAGGTFFEDIAL 223
DB	193 KNFIRPKFNFSKANVLETGNKFLTKCHFS-STNDYCFIFRGLSVIRWAGADFQIALK 251	DB	187 VKNFIRPKFNFSKANVLPITNSTYLTCTRYDKDHPYCPVFLVGVNINWYTFQDLAT 246
QY	225 GGSVGRVHWDCDLDTGDCGCPHYSPF-----QLEKSYNFRITATHWBPQGVYEART 276	QY	224 LGSVGRVHWDCDLDTGDCGCPHYSPFQLEKS-----YNRTATHWBPQGV 273
DB	252 GGVIIGHIHWDCDLDAASHCHPHYFNRLDNKHTQSISSGYNFRARYVRPHGVPRD 311	DB	247 RGSIGIGIENWCDDXDESHCNPEYSTRLDSSSENYTALGPSYGNFRARYYNDAAQOT 306
QY	277 LKLYGIRFDILVTGQAGKFLIPTAVTLTGTAWLGVVTFPCDILLVYVDREAHFYRT 336	QY	274 ARTLLKLYGIRFDILVTGQAGKFLIPTAVTLTGTAWLGVVTFPCDILLVYVDREAHFY 333
DB	312 LKAYRIRFDVINGKAGKFSIPTVINIGSGLMGAGVFPACDMILLYMWSKSFY 366	DB	307 YRNLFKVYGIRFDILVNGKAGKFSIPTVINIGSGLMGAGVFPACDMILLYMWSKSFY 366
QY	337 KYBEAKAPK-----ATANSVMREL 355	QY	334 WRTKYBEAKAPKATANSVMRELALASQAR-----LAECLRSSAPATA----- 377
DB	372 KFEKVRGQKEEDNVEVEANEMEQL 396	DB	367 RETKFEAIKKQSERES--REQERKHARHRRHHHHRRQDGRHREKEKPTAEMQPLTSI 424
RESULT 11			
Q8JFP7	PRELIMINARY; PRT; 481 AA.	Q8JFP7	PRELIMINARY; PRT; 422 AA.
AC	Q8JFP7; 22, Created	AC	Q8JFP7; 23, Created
DT	01-OCT-2002 (T-EMBLrel. 22, Last sequence update)	DT	01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT	01-OCT-2002 (T-EMBLrel. 22, Last sequence update)	DT	01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE	ATP-gated ionotropic receptor subunit 5.	DE	Purinergic receptor P2X5, isoform A.
GN	Name=p2rx5; Synonyms=p2x5;	GN	Name=p2rx5;
OS	Brachydanio rerio (Zebrafish) (Danio rerio).	OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Osteichthyes; Cypriniformes;	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC	Cyprinidae; Danio.	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_taxid=7955;	OX	NCBI_taxid=9606;
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
TI	TISSUE=Whole animal;	TI	TISSUE=Lung;
RX	MEDLINE=22122299; PubMed=12127972;	RX	MEDLINE=22389257; PubMed=12477932;
RA	Diaz-Hernandez M., Cox J.A., Migita K., Haines W., Egan T.M.,	RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Voigt M.M.,	RA	Klausner R.G., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RT	"Cloning and characterization of two novel zebrafish P2X receptor	RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RT	subunits."		
RL	Biochem. Biophys. Res. Commun. 295:849-853(2002).		
CC	-!- SUBUNIT: Homo- or heteropolymers (By similarity).		
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).		
CC	-!- SIMILARITY: Belongs to the P2X receptor family.		



RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	384	TQTPGWPCP	392
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,	405	QKNGSVCP	413
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.,			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RL	SEQUENCE FROM N.A.			
RP	TISSUE=Brain;			
RA	Strausberg R.,			
RA	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
CC	-!- SUBUNIT: Homo- or heteropolymers (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).			
CC	-!- SIMILARITY: Belongs to the P2X receptor family.			
DR	EMBL; BC039015; AAH39015.1; ..			
DR	GO; GO:0016021; C:integral to membrane; IEA.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0005216; F:ion channel activity; IEA.			
DR	GO; GO:0004872; F:receptor activity; IEA.			
DR	GO; GO:0006811; P:ion transport; IEA.			
DR	InterPro; IPR003048; P2X5_receptor.			
DR	InterPro; IPR001429; P2X_receptor.			
DR	Pfam; PF00864; P2X_receptor; 1.			
DR	PRINTS; PR01312; P2XSRECEPTOR.			
DR	PRINTS; PR01307; P2XRECEPTOR.			
DR	TIGRFAMs; TIGR00863; P2X; 1.			
DR	PROSITE; PS01212; P2X RECEPTOR; 1.			
KW	Glycoprotein; Ion transport; Ionic channel; Receptor; Transmembrane;			
KW	Transport.			
SQ	SEQUENCE 422 AA; 47204 MW; B122027A23B2A240 CRC64;			
Query Match 36.9%; Score 822; DB 2; Length 422;				
Best Local Similarity 40.3%; Pred. No. 1.5e-66;				
Matches 173; Conservative 60; Mismatches 120; Indels 76; Gaps 9;				
QY	12 LLDYKTEK-----WALLAKYQERDELPQFSITKLKG 45			
DB	13 LFDYKTEKVIANKKVGILLYLLOASTILAYLVWVFLIKGYQDVTSLOSATVTKVG 72			
QY	46 VSVTQIKELGNLWDVADVPKQENVFVLTNPLVPAQVQRCPEHPSPVLANCWD 105			
DB	73 VAFVTSDLGQINDVADVIVIAQGENVFVTLNLTVPNGQNVCAENEGIPDGACSKD 132			
QY	106 EDCPEGEGTHGHVNTGQCVFNGTHR-TCIWNQCVESGVSPRPLLAQONFTLFI 164			
DB	133 SDCHAGEAVTAGVTKGTCGLRENLAGTCIFAWCPLETSSRDEEFLKEAEDFTFI 192			
QY	165 KNTVTFSEKFNKSALETWDTYFKHCRYPQFSPVFRIGLVAKAGTTEDLALL 224			
DB	193 KKHIRPPKFNKSNVMDVKDSFLKSHFGPK-NHYCPIFLGLSVIWAQSDQDIALE 251			
QY	225 GGSVGIRVHWCDDLTGDSGCPHYXSF-QLQEK-----SYNFRATHWWEQGVFART 276			
DB	252 GGVIGINIEWNCDLKAASECHPHYSFSLDNKLKSVSSGVNFRFARYRDAAGVEPT 311			
QY	277 LKLVGIRFDILVTGQAGKEGLIPAVTLGTGAAMLGVVTFPCDLLLYVDREAHFYWT 336			
DB	312 LMKAYGIRFDVWNGK-----AFFCDLVLYLTKREFYRDK 349			
QY	337 KYEEAKAPKATANSVWRELALASQARLACLRSSAPA-----PTATAAGSQ 383			
DB	350 KYEEVRGLEDSQEADE---ASGLGLSEQL--TSGPGLLGMPEQQLQEPPEAKRGSS 404			

RESULT 13

P2X4\_HUMAN

ID P2X4\_HUMAN STANDARD; PRT; 388 AA.

AC Q9571; O00450; O14722; Q9UBG9;

DT 01-NOV-1997 (Rel. 35, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE P2X purinoceptor 4 (ATP receptor) (P2X4) (Purinergic receptor).

GN Names=P2RX4;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

TISSUE=Brain;

RC MEDLINE=97168759; PubMed=9016352;

RA Garcia-Guzman M., Soto F., Gomez-Hernandez J.M., Lund P.E.,

RA Stuhmer W.;

RT "Characterization of recombinant human P2X4 receptor reveals

RT pharmacological differences to the rat homologue.";

RL Mol. Pharmacol. 51:109-118(1997).

[2]

RP SEQUENCE FROM N.A.

RA Yamagata K., Oda N., Furuta H., Vaxillaire M., Souham L., Boriraj V.,

RA Chen X., Oda Y., Takeda J., Yamada S., Nishigori H., Lebeau M.M.,

RA Lathrop M., Cox R.D., Bell G.I.;

RT "Transcription map of the 5CM region surrounding the hepatocyte

RT nuclear factor-1a/MODY3 gene on chromosome 12.";

RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE FROM N.A.

RA Korenaga R., Yamamoto K., Kamiya A., Ando J.;

RT "Shear stress downregulates the expression of P2X4 receptor by human

RT endothelial cells.";

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

[4]

RP SEQUENCE OF 35-388 FROM N.A.

RA Takahashi K., Korenaga R., Kamiya A., Ando J.;

RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

[5]

RP SEQUENCE OF 73-326 FROM N.A.

TISSUE=Kidney;

RA Chang A.S., Chang S.M.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Receptor for ATP that acts as a ligand gated ion

CC channel. This receptor is insensitive to the antagonists PPADS and

CC suramin.

CC -!- SUBUNIT: Homo- or heteropolymers (By similarity).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: Belongs to the P2X receptor family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; Y07684; CAA68948.1; ..

DR EMBL; U87270; AAD00556.1; ..

DR EMBL; U85971; AAD00556.1; JOINED.

DR EMBL; U85972; AAD00556.1; JOINED.

DR EMBL; U85973; AAD00556.1; JOINED.

DR EMBL; U85974; AAD00556.1; JOINED.

DR EMBL; U85975; AAD00556.1; JOINED.



DR EMBL; U83993; AAD00553.1; --  
 DR EMBL; AF191093; AAR06661.1; --  
 DR EMBL; AF000234; AAB58405.1; --  
 DR EMBL; AF012903; AAB66834.1; --  
 DR Genew; HGNC:8535; P2RX4.  
 DR MIM; 600846; --  
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.  
 DR GO; GO:0004931; P: ATP-gated cation channel activity; TAS.  
 DR GO; GO:007165; P: signal transduction; TAS.  
 DR GO; GO:0006810; P: transport; TAS.  
 DR InterPro; IPR001429; P2X receptor.  
 DR Pfam; PF00864; P2X receptor; 1.  
 DR PRINTS; PRO1307; P2XRECEPTOR.  
 DR TIGRFAMS; TIGR00863; P2X; 1.  
 DR PROSITE; PS01212; P2X\_RECEPTOR; 1.  
 KW Glycoprotein; ion transport; Ionic channel; Polymorphism; Receptor;  
 KW Transmembrane.  
 FT DOMAIN 1 33 Cytoplasmic (Potential).  
 FT TRANSMEM 34 54 1 (Potential).  
 FT DOMAIN 55 338 Extracellular, cysteine-rich (Potential).  
 FT TRANSMEM 339 359 2 (Potential).  
 FT DOMAIN 360 388 Cytoplasmic (Potential).  
 FT CARBOHYD 75 75 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 110 110 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 153 153 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 184 184 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 199 199 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 208 208 N-linked (GlcNAc... ) (Potential).  
 FT VARIANT 242 242 S->G (in dbSNP:25644).  
 FT /FTID=VAR\_014942.  
 FT CONFLICT 6 6 S->A (in Ref. 2).  
 FT CONFLICT 121 121 D->S (in Ref. 5).  
 FT CONFLICT 148 148 R->W (in Ref. 5).  
 FT CONFLICT 179 179 L->P (in Ref. 5).  
 FT CONFLICT 238 238 N->I (in Ref. 5).  
 FT CONFLICT 315 315 Y->C (in Ref. 4).  
 SQ SEQUENCE 388 AA; 43385 MW; 3681E192E9CF201E CRC64;  
 Query Match 36.7%; Score 816; DB 1; Length 388;  
 Best Local Similarity 47.0%; Pred. No. 4.7e-66;  
 Matches 155; Conservative 56; Mismatches 107; Indels 12; Gaps 4;  
 QY 20 WALLAKKGQERDLEPQFSIIITKLKGVSVTQIKELGNRLWDVADVPKPGQGNVFLVTN 79  
 Db 45 WYFVMEKGYQETD-SVYSSVTTTKVGAVTNTSKLGFRIWDVADYVIPAQEENSFLVMTN 104  
 QY 80 FLVTPAQVQGRCPHPSPVLANCWVDEDCPEGGTHSHGVKTGCQVFNTHRTCEIWS 139  
 Db 105 VILTMNQGLCPETDATTV-CKSDASCTAGSAGTHNGVSTGRCAVNGSVKTCVAA 163  
 QY 140 WCPVSEGV-VPSRPLLAQAQNTFLIKNTVTSKFNFSKNALETWDPDYFKHCRYPEQF 198  
 Db 164 WCPVEDDTHVPQAPFLKAAENFTLLVKNWYKFNFSKRNILPNITTYLKSCIYDAKT 223  
 QY 199 SPYCPVFRIGDLVAKGTFEDLALLGSGVGRVHMDCLDTGDSGCPHYSPQLOE--- 255  
 Db 224 DPFCIFRLCKIVENAGHSFQDAVEGGIMGIVQNVDCNLDRAASLCPLRYSFRRLDTRD 283  
 QY 256 -----KSNYFRTATHWEEQVGEARTLLKLYGIRFDILVTGQAGKFGLIPTAVTLGTA 309  
 Db 284 VEHNVSPGYNFRFAKYRDLACNEQRTLLIKAYGIRFDIIVFGKAGKFDIITMINISGL 343  
 QY 310 ANLGVVTFPCDLLLVYDREAHFYRTKYE 339  
 Db 344 ALLGMATVLCIDIIVLYCMKRLYYREKRYK 373  
 RESULT 14  
 Q8N4N1 PRELIMINARY; PRT; 388 AA.  
 ID Q8N4N1  
 AC Q8N4N1;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT EMBL; U83993 (TrEMBLrel. 25, Last annotation update)  
 DE Purinergic receptor P2X4, isoform a.  
 GN Name=P2RX4;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA Strausberg R.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBUNIT: Homo- or heteropolymers (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to the P2X receptor family.  
 DR EMBL; BC033826; AAB33826.1; --  
 DR PIR; JC6543; JC6543.  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR GO; GO:0005524; F: ATP binding; IEA.  
 DR GO; GO:0005216; F: ion channel activity; IEA.  
 DR GO; GO:0004872; P: receptor activity; IEA.  
 DR GO; GO:0006811; P: ion transport; IEA.  
 DR InterPro; IPR003047; P2X4\_purinceptor.  
 DR InterPro; IPR001429; P2X\_receptor; 1.  
 DR Pfam; PF00864; P2X\_receptor; 1.  
 DR PRINTS; PRO1311; P2X4RECEPTOR.  
 DR PRINTS; PRO1307; P2XRECEPTOR.  
 DR TIGRFAMS; TIGR00863; P2X; 1.  
 DR PROSITE; PS01212; P2X\_RECEPTOR; 1.  
 KW Glycoprotein; ion transport; Ionic channel; Receptor; Transmembrane;  
 KW Transport.  
 SQ SEQUENCE 388 AA; 43339 MW; AF7ABC688FCE02A5 CRC64;  
 Query Match 36.5%; Score 813; DB 2; Length 388;  
 Best Local Similarity 47.0%; Pred. No. 8.8e-66;  
 Matches 155; Conservative 55; Mismatches 108; Indels 12; Gaps 4;  
 QY 20 WALLAKKGQERDLEPQFSIIITKLKGVSVTQIKELGNRLWDVADVPKPGQGNVFLVTN 79  
 Db 46 WYFVMEKGYQETD-SVYSSVTTTKVGAVTNTSKLGFRIWDVADYVIPAQEENSFLVMTN 104  
 QY 80 FLVTPAQVQGRCPHPSPVLANCWVDEDCPEGGTHSHGVKTGCQVFNTHRTCEIWS 139  
 Db 105 VILTMNQGLCPETDATTV-CKSDASCTAGSAGTHNGVSTGRCAVNGSVKTCVAA 163  
 QY 140 WCPVSEGV-VPSRPLLAQAQNTFLIKNTVTSKFNFSKNALETWDPDYFKHCRYPEQF 198  
 Db 164 WCPVEDDTHVPQAPFLKAAENFTLLVKNWYKFNFSKRNILPNITTYLKSCIYDAKT 223



Db 349 KYEEVRGLEDSSQAEDE---ASGLGLSEQL--TSGPGLLGMPQQELQEPPEAKRGSSS 403

Qy 384 TQTEGWPCP 392

Db 404 QKNGSVCP 412

Search completed: November 21, 2004, 15:03:29  
Job time : 193 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:54:02 ; Search time 40 Seconds  
(Without alignments)  
974.195 Million cell updates/sec

Title: US-09-820-095B-2  
Perfect score: 2226  
Sequence: 1 MGSPGATGWLDDYKTEK.....TPGWPCSSDTHLPTHSGL 405

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1554	69.8	379	JC4843	ATP-gated ionchann
2	900	40.4	455	S71344	purinergic recepto
3	815	36.6	388	JC6543	P2X4 receptor - hu
4	793	35.6	388	JC6137	ATP-gated ion chan
5	744	33.4	399	S71327	ATP-gated ion chan
6	733	33.2	339	S50860	apoptosis protein
7	720	32.3	472	S50859	P2x receptor - rat
8	679.5	30.5	397	S60334	purinoceptor P2X -
9	679.5	30.5	397	I58099	gene P2X3 protein
10	96.5	4.3	562	B91194	probable enzyme (i
11	96.5	4.3	562	C86041	probable enzyme yi
12	95	4.3	372	T47344	hypothetical prote
13	95	4.3	1036	F71565	probable isoleucin
14	94.5	4.2	562	A65166	hypothetical 63.2K
15	94.5	4.2	628	I38000	Lutheran blood gro
16	92.5	4.2	588	I37202	B-CAM protein - hu
17	91.5	4.1	392	I54491	cell surface antig
18	90	4.0	844	C91212	probable outer mem
19	90	4.0	844	C86058	probable fibrial
20	89.5	4.0	1025	A43526	complement C3d/Eps
21	88	4.0	3947	T52486	ferrichrome sidero
22	87	3.9	6658	T13931	projectin - fruit
23	86.5	3.9	618	T00476	probable vacuolar
24	86.5	3.9	737	S65758	nitrate reductase
25	86	3.9	375	A32827	fetuin precursor -
26	85.5	3.8	1743	T28859	hypothetical prote
27	85	3.8	806	A35963	protein-tyrosine k
28	84.5	3.8	282	JC5439	D-aspartate oxidas
29	84.5	3.8	341	JC5438	D-aspartate oxidas

30	84.5	3.8	559	1	A35029	t-plasminogen acti
31	84.5	3.8	995	2	C83203	probable serine pr
32	84	3.8	561	2	A25090	chitinase (EC 3.2
33	83.5	3.8	635	2	A45266	MPL-P protein prec
34	83.5	3.8	1036	2	F81719	isoleucyl-tRNA syn
35	83.5	3.8	1367	2	T21913	hypothetical prote
36	83	3.7	347	2	T08371	probable UDP-glucu
37	83	3.7	364	2	AE3196	conserved hypothet
38	83	3.7	713	1	S30019	acetate-CoA ligase
39	83	3.7	881	2	AD2580	two component sens
40	83	3.7	881	2	C97362	protein sensor pro
41	83	3.7	1049	2	T30525	alpha-mannosidase
42	82.5	3.7	164	2	H95889	phosphoesterase-re
43	82.5	3.7	508	2	S41879	readthrough protei
44	82.5	3.7	746	2	AC1883	nitrate reductase
45	82.5	3.7	1344	2	S47412	gene P2 protein -

ALIGNMENTS

RESULT 1

JC4843

ATP-gated ionchannel receptor, P2X6 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 15-Aug-1996 #sequence\_revision 15-Oct-1996 #text\_change 09-Jul-2004

C:Accession: JC4843

R:Soto, F.; Garcia-Guzman, M.; Karschin, C.; Stuehmer, W.

Biochem. Biophys. Res. Commun. 223, 456-460, 1996

A:Title: Cloning and tissue distribution of a novel P2X receptor from rat brain.

A:Reference number: JC4843; MUID:96264662; PMID:8670303

A:Contents: brain

A:Accession: JC4843

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-379 <SOT>

A:Cross-references: UNIPROT:P51579; GB:X97376; MID:g1419751; PIDN:CAA66044.1; PID:g14197

C:Comment: This receptor is a ATP-gated ion channel.

C:Superfamily: ATP-gated ion channel P2X4 receptor

C:Keywords: brain; glycoprotein; receptor; transmembrane protein

F:30-52/Domain: transmembrane #status predicted <TM1>

F:330-355/Domain: transmembrane #status predicted <TM2>

F:157,187,202/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	69.8%	Score	1554	DB 2	Length	379			
Best Local Similarity	75.7%	Pred. No.	7.2e-130						
Matches	281	Conservative	26	Mismatches	38	Indels	26	Gaps	1
QY	6	ATTGGLLDYKTEK-----	WALLAKGYQERDLEPQFSI	39					
DB	8	ALVSWGFLDYKTEKYVMTRNCWVGISQRLQLGVVVVIGWALLAKGYQEWMDPQISV	67						
QY	40	ITKLKGVSVTQIKELGNLMDVADVPKPPQGENVFFLVTNFLVTPAQVQGRCPHSPVPL	99						
DB	68	ITKLKGVSVTVQKELEKRLMDVADPVRPSQGENVFFLVTNFLVTPAQVQGRCPHSPVPL	127						
QY	100	ANCWVDEDCPGEGETHSHGVKTCQCQVFNQTHRTCTEIWSMCPVESGVVPRPPLAAQKN	159						
DB	128	ANCWADEDCPGEEMGTYSHGKTCQCVAFNGTHRTCTEIWSMCPVESSAVPRPKPILAAQKN	187						
QY	160	FTLFKNTVTPFSKNFNSKNALETWDTPYFKHCRIYEPQFSPYCPVFRIGDLVAKAGTFPE	219						
DB	188	FTLFKNTVTPFNKFNFTNALTNDTNTYFKCYDLSLSSPSPVFRIGDLVAMTGGDFE	247						
QY	220	DLALLGSGVIRVHWDCCDLDTGDSGCWPHYSFQLOEKSYNFRATTHWEGPVEARTLLK	279						
DB	248	DLALLGSAVGINIIHWDCNLDTKSDCSFQSFQLOEGYNFRTANYWASGVESRILK	307						
QY	280	LYGIRFDILVTQAGKFGLIPTAVTLTGTGAALGWVTFPCDLLLLYYDRHAHFVWRRTKYE	339						
DB	308	LYGIRFDILVTQAGKFALIPTAITVGTGAALGWVTFPCDLLLLYYDRAGFWRTKYE	367						
QY	340	EAKAPKATANS	350						

Db 368 EAPAPKATNS 378  
 ||:|||||

## RESULT 2

S71344  
 purinergic receptor P2X5 - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 09-Jul-2004  
 C/Accession: S71344  
 R/Garcia-Guzman, M.; Soto, F.; Laube, B.; Stuehme, W.  
 FEBS Lett. 388, 123-127, 1996  
 A/Title: Molecular cloning and functional expression of a novel rat heart P2X purinoceptor  
 A/Reference number: S71344, MUID:96275514, PMID:8690069  
 A/Accession: S71344  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-455 <GAR>  
 A/Cross-references: UNIPROT:P51578; EMBL:X937328; NID:g1419749; PIDN:CAA65993.1; PID:g1419749  
 A/Experimental source: heart  
 C/Genetics:  
 A/Gene: P2X5  
 C/Superfamily: ATP-gated ion channel P2X4 receptor  
 C/Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein  
 F:29-51/Domain: transmembrane #status predicted <TM1>  
 F:337-362/Domain: transmembrane #status predicted <TM2>  
 F:18,149/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F:26,158/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted  
 F:77,157,202/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 40.4%; Score 900; DB 2; Length 455;  
 Best Local Similarity 46.9%; Pred. No. 8.9e-72; Indels 36; Gaps 5;  
 Matches 175; Conservative 55; Mismatches 107;

QY 12 LLDYKTEK-----WALLAKGYQERDLEPPQPSIITLKG 45  
 Db 13 LPDYKAKFVAKSKVGLLYRVLQILLLYLLVFLIKSYQIDITSLQSAVTVKVG 72  
 QY 46 VSVTOIKELGNLWDVADFKVPPQGENVFLVNTFLVTPAQVQGRCPHSPVPLANCWD 105  
 Db 73 VAYNTTMLGERLWDVADFKVPPQGENVFLVNTFLVTPAQVQGRCPHSPVPLANCWD 132  
 QY 106 EDCPEGGTHSHGVTKGCV-VFNGTHTCEIWSMCPVSGVPSRPLLAQAQNTLFI 164  
 Db 133 DDCHAGESVAGGLTKGCLRNSTRGTCEIFAWCPVETKSMPTDPLKDAESFTFI 192  
 QY 165 KNTVTFKSNFSSNALETWDPTFKHCRYEPQSPYCPVFRIGDLVAKAGTFFEDLALL 224  
 Db 193 KNFIRPKFNFSKANVLETDNKHLXTCHFS-STNLYCPIFLGSIWRWAGADFQDIALK 251  
 QY 225 GGSVGRVHWDCDLTDGSCWPHYSF-QLQEK-----SVNERTATHWQPGVEART 276  
 Db 252 GGVIGVIEWDCDLDAASKNCPHYFNRLDNKHTSISSGVNFRFARYRDPNGVEFRD 311  
 QY 277 LLKLYGIRPDILVTGQAGKFGLLIPTAVLTGTGAALVGVVTFPCDLLLIVYDREAHFYRT 336  
 Db 312 LMKAYGIRFDVTVNGKAGKFSIPTVINIGSGLALGAGAFFCDLVLILRKSEFYRDK 371  
 QY 337 KYEAKAPKATAN 349  
 Db 372 KFEKVRGQKEDAN 384

## RESULT 3

JC6543  
 P2X4 receptor - human  
 C/Species: Homo sapiens (man)  
 C/Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 09-Jul-2004  
 C/Accession: JC6543  
 R/Dhulipala, P. D. K.; Wang, Y. X.; Kotlikoff, M. I.  
 Gene 207, 259-266, 1998  
 A/Title: The human P2X4 receptor gene is alternatively spliced.  
 A/Reference number: JC6543; MUID:98172760; PMID:9511769

A/Accession: JC6543  
 A/Molecule type: mRNA  
 A/Residues: 1-388 <DHU>  
 A/Cross-references: UNIPROT:Q8N4N1  
 A/Experimental source: stomach  
 C/Comment: This receptor is an ATP requiring ion channel. It mediates membrane depolariza  
 C/Genetics:  
 A/Introns: 95/3  
 C/Superfamily: ATP-gated ion channel P2X4 receptor  
 C/Keywords: glycoprotein; ion channel; receptor  
 F:30-50/Domain: transmembrane #status predicted <TM1>  
 F:338-361/Domain: transmembrane #status predicted <TM2>  
 F:75,110,153,184,199,287/Binding site: carboxylate (Asn) (covalent) #status predicted  
 Query Match 36.6%; Score 815; DB 2; Length 388;  
 Best Local Similarity 47.0%; Pred. No. 2.5e-64;  
 Matches 155; Conservative 56; Mismatches 107; Indels 12; Gaps 4;  
 QY 20 WALLAKGYQERDLEPPQPSIITLKGYSVTQIKELGNRLWDVADFKVPPQGENVFLVNT 79  
 Db 46 WVFVWEKGYQETD-SVYSSVTTKVKGAVTNTSKLGRFDVADYVIPAQKENSFLVMTN 104  
 QY 80 FLVTPAQVQGRCPHSPVPLANCWDDECPGEGGTHSHGVTKGQCVWFNGTHRTCEIWS 139  
 Db 105 VILTNQVQGLCEIPDATTV-CKSPASCTAGSAGTHSGVSTGRCVAFNGSVKTCVAA 163  
 QY 140 WCPVESGV-VPSRPLLAQAQNTLFIKNTVTFKSNFSSNALETWDPTFKHCRYEPQF 198  
 Db 164 WCPVEDDTHVQPAFLKAAENFTLLVKNINWIYKFNKFNILNTTTTLKSCIIYDAKT 223  
 QY 199 SPYCPVFRIGDLVAKAGTFFEDLALLGSGVGRVHWDCDLTDGSCWPHYSFQLOE--- 255  
 Db 224 DPCPIPRLGIKVIENAGHSFQDMVAVKGGINGIQVWDCNLDRAASLCLPRYSFRLEDTRD 283  
 QY 256 -----KSNFRFATHWQPGVEARTLLKLYGIRPDILVTGQAGKFGLLIPTAVLTGCA 309  
 Db 284 VEHNVSPGVNFRFARYRDLAGNEQRTLIKAYGIREDFIVFGKAGKFDIPTMINIGSGL 343  
 QY 310 AWLGVTFFCDLLLVYDREAHFYRTKYE 339  
 Db 344 ALLGMSTVLCIDIIVLYCMKRLIYREKKYK 373  
 RESULT 4  
 JC6137  
 ATP-gated ion channel P2X4 - rat  
 N/Alternate names: ATP-activated ionotropic receptor P2X4; P2X4 purinoreceptor  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004  
 C/Accession: JC6137; JC4684; S6206; S72563; S62359  
 R/Soto, F.; Garcia-Guzman, M.; Gomez-Hernandez, J. M.; Hollmann, M.; Karschin, C.; Stuehme  
 Proc. Natl. Acad. Sci. U.S.A. 93, 3684-3688, 1996  
 A/Title: P2X4: An ATP-activated ionotropic receptor cloned from rat brain.  
 A/Reference number: JC6137; MUID:96195033; PMID:8622997  
 A/Contents: brain  
 A/Accession: JC6137  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-388 <SOT>  
 A/Cross-references: UNIPROT:P51577; EMBL:X93565; NID:g1325958; PIDN:CAA63778.1; PID:g1325958  
 R/Wang, C. Z.; Namba, N.; Gonoi, T.; Inagaki, N.; Seino, S.  
 Biochem. Biophys. Res. Commun. 220, 196-202, 1996  
 A/Title: Cloning and pharmacological characterization of a fourth P2X receptor subtype w  
 A/Reference number: JC4684; MUID:96176973; PMID:8602843  
 A/Accession: JC4684  
 A/Molecule type: mRNA  
 A/Residues: 1-388 <GB>  
 A/Cross-references: UNIPROT:U47031; NID:g1184996; PIDN:AAA99777.1; PID:g1184997  
 A/Experimental source: pancreatic islet  
 R/Bo, X.; Zhang, Y.; Nassar, M.; Burnstock, G.; Schoeffer, R.  
 FEBS Lett. 375, 129-133, 1995  
 A/Title: A P2X purinoceptor cDNA conferring a novel pharmacological profile.







QY 138 WSWCPVSVGVPSPRLLAQAOQNTFLFKNTVTFSKFNFSKNALETWDPYFKHCRYPQP 197  
Db 150 QGWCPEVDTV-EMPIMAEAFNTIFIKNSIRPLNFKGNLLNLTDKDIKRCRPHPE 208  
QY 198 FSPYCPVFRIGDLVAKAGTFFDLALGSGVIRVHWCDDLDTGSGCWPVHSPFO----L 253  
Db 209 KAPFCPIELVGVGVFAGQDFAKLARTGCVLGKIGWVCDLKAWDQCPKYSFTRLDGV 268  
QY 254 QSKS-----YNFRTATHWQVQVEARTLLKLYGIRFDILTQVQAGKFLIPTAVTLGTG 308  
Db 269 SEKSVSPGPNFRFAKYKXGSEYRULLKAFGRFVLYVGNAGKNLTIPTISSVAA 328  
QY 309 AAWLGVTVPFCDLLLLYVDREAHFVWRTKYBEAKAP--KATANS 350  
Db 329 FTSVGVTGLCDIILLNFKLGADHYKARKFEETVETLLKGTAST 372

RESULT 10  
B91194  
probable enzyme [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: B91194  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic islands  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: B91194  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-562 <HAY>  
A:Cross-references: UNIPROT:Q8XD89; GB:BA000007; PIDN:BA037945.1; PID:G13363997; GSPDB:G13363997  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECs4522

A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: C86041  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-562 <STO>  
A:Cross-references: UNIPROT:Q8XD89; GB:AE005174; NID:q12518406; PIDN:AAG58791.1; GSPDB:G13363997  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: yicP

Query Match 4.3%; Score 96.5; DB 2; Length 562;  
Best Local Similarity 18.6%; Pred. No. 2.1;  
Matches 50; Conservative 40; Mismatches 104; Indels 75; Gaps 9;  
QY 14 DYKTEKALLAKGYQERDLEPQFSIITKLKGVSVTQIKELGNRLNDVADVPKPGQGNV 73  
Db 342 DKXQVRNIGSVRRWQEWDIAPGDQILVSLAQGIPRIDDV--VVRGAERTKPTPPENR 398  
QY 74 FFLVTNFLTVAQVQGRCEPHSPVPLANCWVDEDCPEGEGTHSHGVKTKQCVVFNQTHR 133  
Db 399 FNSLTCTCFASDV-----CQQFISRLV--WLGSKQVLGLDGIAGEAGWR-----ALHQTHR 446  
QY 134 TCEIWSWCPVSVGVPSPRLLAQAOQNTFLFKNTVTFSKFNFSKNALETWDPYFKHCR 193  
Db 447 FEHIFSWLLLTPEQLQNTPGIAKS-----KSAQLWHRFNLRKQKQPTTRW-----490  
QY 194 YEPQSPYCPVFRIGDLVAKAGTFFDLALGSGVIRVHWCDDLDTGSGCWPVHSPFOL 253  
Db 491 -----VMANGIPLTRAA-----LNASDERSWSQLLFS- 517  
QY 254 QKSYNFRATATHWQVQGV---EARTLLK 279  
Db 518 -----TEQFWQQLFGTSGRGRQVIE 538

RESULT 12  
T47344  
hypothetical protein F18P9.20 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T47344  
R:Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24458  
A:Accession: T47344  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-372 <NYA>  
A:Cross-references: UNIPROT:Q9M1L9; EMBL:AL138654  
A:Experimental source: cultivar Columbia; BAC clone F18P9  
C:Genetics:  
A:Map position: 3  
A:Introns: 125/3; 194/3  
A>Note: F18P9.20

Query Match 4.3%; Score 95; DB 2; Length 372;  
Best Local Similarity 21.7%; Pred. No. 1.6;  
Matches 73; Conservative 33; Mismatches 110; Indels 120; Gaps 20;  
QY 16 KTERKALLAKGYQERDLEPQFSIITKLKGVSVT-----QIKELGNRLNDVADVPK--PQ 69  
Db 51 KSDQW-----CQSELPN-----ASKRSVAVTTGGFQSRDGGGVAGQDFEKSQPC 98  
QY 70 GENVFELVT-NFLVTPAQVQGRCEPH-----PSVPLANCW-----103  
Db 99 GVGICLILTSNTPKNPKYKPNRENGCGFFQWCDAVQSGTGTSTTSNSYNGNDT 158  
QY 104 --VDEDCPEGEGTHSHGVKTKGQ-----CVVFNGTHRTCEIWSWCPVSVGVPSPRPL 153  
Db 159 KFPDHCQPCGAGLCRVLTAKTGENVGRQYRCPVFEK---SCGFKWC--NDNVSSPTS 213  
QY 154 LAQAQNTFLFKNTVTFSKFNFSKNA-----LSTWDPYFKHCRYPQFSPYCPV----204

QY 138 WSWCPVSVGVPSPRLLAQAOQNTFLFKNTVTFSKFNFSKNALETWDPYFKHCRYPQP 197  
Db 150 QGWCPEVDTV-EMPIMAEAFNTIFIKNSIRPLNFKGNLLNLTDKDIKRCRPHPE 208  
QY 198 FSPYCPVFRIGDLVAKAGTFFDLALGSGVIRVHWCDDLDTGSGCWPVHSPFO----L 253  
Db 209 KAPFCPIELVGVGVFAGQDFAKLARTGCVLGKIGWVCDLKAWDQCPKYSFTRLDGV 268  
QY 254 QSKS-----YNFRTATHWQVQVEARTLLKLYGIRFDILTQVQAGKFLIPTAVTLGTG 308  
Db 269 SEKSVSPGPNFRFAKYKXGSEYRULLKAFGRFVLYVGNAGKNLTIPTISSVAA 328  
QY 309 AAWLGVTVPFCDLLLLYVDREAHFVWRTKYBEAKAP--KATANS 350  
Db 329 FTSVGVTGLCDIILLNFKLGADHYKARKFEETVETLLKGTAST 372

RESULT 11  
C86041  
probable enzyme yicP [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: C86041  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.



```
QY 50 QIKELGNRL--WDVADFKPPQGENVFLVTNPLVTPAQVQGRCPHP--SVPLANC--- 102
Db 331 QSGTYGCRVEDYDAADDV---QLSKTLELRVAYLDPLELSEKVLSLPLNSSAVVNCVSH 387
QY 103 -----WDEDCPEGEGTHSHGVKTCQCVF--NGHRTCEIWSWCPVESGVVPSRP 152
Db 388 GLFPTALRWTKDSTPLGDGMLSL-----SSITFDSNGTY-VCE-----ASLPTVP 432
QY 153 LLAQAQNFTLFIKNTVTFSKFNFSKNALETWDPYFKHCYBPQFSPYCPVFRIGDLVA 212
Db 433 VLSRTQNFLLVQGS-----PELKTAEIEP 457
QY 213 KAGGFEDLALGGSVGI-----RVHWCDDLDGDSG-----CWPHYSFOLQKS 257
Db 458 KADGSWRE---GDEVTLICSARGHPDPKLSWSQLGGSFAEPIPGRQGWVSSSLTK--- 510
QY 258 YNFRTATHWEQPGVEARTL---LKLXGIRFDILVTGQAGKFGJLPTAVTLGTGAAMLG 313
Db 511 -----VTSALSRDGISCEASNPHGNKRHFHFG-AVSPQTSQAGVAVMAVSVGLULLLV 564
QY 314 VVTFFC 319
Db 565 VAVFYC 570
```

Search completed: November 21, 2004, 15:04:20  
Job time : 47 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: November 21, 2004, 15:00:26 ; Search time 144 Seconds  
(without alignments)  
995.984 Million cell updates/sec

Title: US-09-820-095B-2

Perfect score: 2226

Sequence: 1 MGSPATTGGLDYKTEK.....TPGWPCSSDTHLTHSGSL 405

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2226	100.0	405	10	US-09-820-095-2
2	2203	99.0	431	10	US-09-820-095-4
3	1080.5	48.5	364	15	US-10-051-874-125
4	825.5	37.1	422	17	US-10-128-558-144
5	816	36.7	388	9	US-09-833-082-2
6	816	36.7	388	15	US-10-455-552-2
7	810	36.4	388	15	US-10-386-414-17
8	744	33.4	399	14	US-10-352-684A-54
9	709	31.9	459	14	US-10-345-680-11
10	709	31.9	459	15	US-10-051-874-123
11	704.5	31.6	397	16	US-10-408-765A-2202
12	694	31.2	402	11	US-09-764-875-905
13	694	31.2	404	15	US-10-051-874-124

14	693	31.1	497	15	US-10-051-874-120	Sequence 120, App
15	673.5	30.3	287	15	US-10-455-552-3	Sequence 3, Appl1
16	631	28.3	247	15	US-10-051-874-121	Sequence 121, App
17	631	28.3	447	15	US-10-051-874-122	Sequence 122, App
18	615	27.6	473	15	US-10-051-874-42	Sequence 42, Appl
19	611.5	27.5	595	16	US-10-408-765A-2166	Sequence 2166, Ap
20	608.5	27.3	595	16	US-10-622-313-1	Sequence 1, Appl1
21	608.5	27.3	595	16	US-10-789-241-40	Sequence 40, Appl1
22	606.5	27.2	595	10	US-09-977-221-4	Sequence 4, Appl1
23	606.5	27.2	595	16	US-10-766-978-4	Sequence 4, Appl1
24	583	26.2	348	15	US-10-336-472-52	Sequence 52, Appl1
25	545.5	24.5	274	15	US-10-336-472-54	Sequence 54, Appl1
26	500	22.5	260	14	US-10-104-047-2151	Sequence 2151, Ap
27	379.5	17.0	280	9	US-09-764-847-739	Sequence 739, App
28	379.5	17.0	280	11	US-09-764-875-1131	Sequence 1181, Ap
29	373.5	17.0	280	14	US-10-092-154-739	Sequence 739, App
30	250	11.2	50	9	US-09-864-761-38753	Sequence 38753, A
31	206	9.3	46	9	US-09-864-761-35496	Sequence 35496, A
32	102	4.6	636	14	US-10-156-761-13432	Sequence 13432, A
33	98.5	4.4	442	15	US-10-425-114-58714	Sequence 58714, A
34	98.5	4.4	450	17	US-10-425-115-346201	Sequence 346201,
35	98.5	4.4	3564	15	US-10-016-248-45	Sequence 45, Appl1
36	95	4.3	754	15	US-10-108-260A-4450	Sequence 4450, Ap
37	94.5	4.2	394	9	US-09-925-301-1388	Sequence 1388, Ap
38	94.5	4.2	394	15	US-10-264-049-3011	Sequence 3011, Ap
39	94.5	4.2	628	14	US-10-176-847-108	Sequence 108, App
40	94.5	4.2	628	14	US-10-205-823-246	Sequence 246, App
41	94.5	4.2	628	14	US-10-257-021-6	Sequence 6, Appl1
42	94.5	4.2	628	14	US-10-257-021-70	Sequence 70, Appl1
43	94	4.2	1000	10	US-09-823-187-6	Sequence 6, Appl1
44	93.5	4.2	595	15	US-10-276-774-2369	Sequence 2369, Ap
45	93.5	4.2	2669	15	US-10-016-248-4	Sequence 4, Appl1

## ALIGNMENTS

### RESULT 1

US-09-820-095-2  
; Sequence 2, Application US/09820095  
; Publication No. US20030233668A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001202  
; CURRENT APPLICATION NUMBER: US/09/820,095  
; CURRENT FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Human  
US-09-820-095-2

Query Match 100.0%; Score 2226; DB 10; Length 405;  
Best Local Similarity 100.0%; Pred. No. 9.3e-214;  
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGSPATTGGLDYKTEKWKALLAKKGYQERDLEPOPSITTKLKGVSVTQIKELGNLWD	60
Db	1	MGSPATTGGLDYKTEKWKALLAKKGYQERDLEPOPSITTKLKGVSVTQIKELGNLWD	60
Qy	61	VADPVKPPQGENVFFLTNLFVTPAQVQGRCPHPSPVPLANCWVDEDCPEGEGTHSHGV	120
Db	61	VADPVKPPQGENVFFLTNLFVTPAQVQGRCPHPSPVPLANCWVDEDCPEGEGTHSHGV	120
Qy	121	KTGCVVFNHGTHTCEIWSWCPVSGVVPSPRLLAQONFTLPIKNTVTSKFNFSKNA	180
Db	121	KTGCVVFNHGTHTCEIWSWCPVSGVVPSPRLLAQONFTLPIKNTVTSKFNFSKNA	180

181 LETWPTVFKHCRVPEPQSPYCPVPRIGDLVAKAGGTFFEDLALGGSGVIRVHWDCLDT 240  
181 LETWPTVFKHCRVPEPQSPYCPVPRIGDLVAKAGGTFFEDLALGGSGVIRVHWDCLDT 240  
241 GDSGCPHYSFOLQSKSYNFRATATWHPQGVARTLLKLYGIRFDILVTQAGKFGILP 300  
241 GDSGCPHYSFOLQSKSYNFRATATWHPQGVARTLLKLYGIRFDILVTQAGKFGILP 300  
301 TAVTLGTGAAMLGVVTFCDLLLVVDREAHFYRTKYEBKAPKATANSVWRELALASQ 360  
301 TAVTLGTGAAMLGVVTFCDLLLVVDREAHFYRTKYEBKAPKATANSVWRELALASQ 360  
361 ARLAELRRSSAPATAAGSQTQPCWPCSSDTHLPTHSGSL 405  
361 ARLAELRRSSAPATAAGSQTQPCWPCSSDTHLPTHSGSL 405

RESULT 2  
US-09-820-095-4  
; Sequence 4, Application US/09820095  
; Publication No. US20030233668A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001202  
; CURRENT APPLICATION NUMBER: US/09/820,095  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Human  
US-09-820-095-4

Query Match 99.0%; Score 2203; DB 10; Length 431;  
Best Local Similarity 94.0%; Fred. No. 2e-211;  
Matches 405; Conservative 0; Mismatches 0; Indels 26; Gaps 1;

1 MGSPGATTGWLGLDYKTEK-----WALLAKKGQERDLE 34  
1 MGSPGATTGWLGLDYKTEKYVTRNRYGALQRLIQGIVVVVGVWALLAKKGQERDLE 60  
35 POFSIITKLKGVSVTQIKELGNRLNDVADFKPPQGVNFVFLVFNFLVTPAQVQGRCEH 94  
61 POFSIITKLKGVSVTQIKELGNRLNDVADFKPPQGVNFVFLVFNFLVTPAQVQGRCEH 120  
95 PSVPLANCWVDEDCPEGGGTHSHGVKTGCQVFNHGTTCETIWSQCVESGVVPSRPLL 154  
121 PSVPLANCWVDEDCPEGGGTHSHGVKTGCQVFNHGTTCETIWSQCVESGVVPSRPLL 180  
155 AQAQNTLFIKNTVTSKFNFSKNALETWPTVFKHCRVPEPQSPYCPVPRIGDLVAKA 214  
181 AQAQNTLFIKNTVTSKFNFSKNALETWPTVFKHCRVPEPQSPYCPVPRIGDLVAKA 240  
215 GGTFFEDLALGGSGVIRVHWDCLDTGDSGCPHYSFOLQSKSYNFRATATWHPQGV 274  
241 GGTFFEDLALGGSGVIRVHWDCLDTGDSGCPHYSFOLQSKSYNFRATATWHPQGV 300  
275 RTLLKLYGIRFDILVTQAGKFGILPTAVTLGTGAAMLGVVTFCDLLLVVDREAHFYH 334  
301 RTLLKLYGIRFDILVTQAGKFGILPTAVTLGTGAAMLGVVTFCDLLLVVDREAHFYH 360  
335 RTKYEBKAPKATANSVWRELALASQARLAELRRSSAPATAAGSQTQPCWPCSS 394  
361 RTKYEBKAPKATANSVWRELALASQARLAELRRSSAPATAAGSQTQPCWPCSS 420  
395 DTHLPTHSGSL 405  
421 DTHLPTHSGSL 431

RESULT 3  
US-10-051-874-125  
; Sequence 125, Application US/10051874  
; Publication No. US20040005557A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Colman, Steven D  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Boldog, Ferenc AM  
; APPLICANT: Vermet, Corine AM  
; APPLICANT: Li, Li  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Guo, Xiaojia Sasha  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: MacDougall, John R  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Pena, Carol EA  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Miller, Charles E  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Smithson, Glenn D  
; APPLICANT: Baumgartner, Jason C  
; APPLICANT: Herrman, John L  
; APPLICANT: Peyman, John A  
; APPLICANT: Gorman, Linda  
; APPLICANT: Mezes, Peter D  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Grosse, William M  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Rothenberg, Mark  
; APPLICANT: Stone, David J  
; APPLICANT: Burgess, Catherine E  
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF  
; TITLE OF INVENTION: USING THE SAME  
; FILE REFERENCE: 21402-245  
; CURRENT APPLICATION NUMBER: US/10/051,874  
; CURRENT FILING DATE: 2002-09-25  
; PRIOR APPLICATION NUMBER: 60/268,595  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/325,306  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/262,587  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/272,409  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/262,454  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/276,777  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/291,672  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/330,336  
; PRIOR FILING DATE: 2001-10-18  
; PRIOR APPLICATION NUMBER: 60/265,530  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/261,376  
; PRIOR FILING DATE: 2001-01-16  
; NUMBER OF SEQ ID NOS: 269  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 125  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

```
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: P2X_receptor
; OTHER INFORMATION: domain sequence
US-10-051-874-125

Query Match      48.5%; Score 1080.5; DB 15; Length 364;
Best Local Similarity 56.1%; Pred. No. 4.1e-99;
Matches 203; Conservative 44; Mismatches 80; Indels 35; Gaps 4;

QY 14 DYKTEK-----WALLAKKGQERDLEPOFSIITKLKGS 47
Db 2 DYKTPKVVVRNKKVGLLNRLVQLLLVYVVGWVFLIEKGYQDSOTSLQSSVITKVGVA 61

QY 48 VTOIKELGNRLMDVADVPKPOGENVFFLVNPLVTPAQVQGRCPHPSPVLANCWVDED 107
Db 62 VNTSELGRVMDVADYVPPQGENVFFVVTNPIVTPNQTQCEHPHVPDGTCKSDSD 121

QY 108 CPPEGGTHSHGVKTCQCVVNGT-HRTCEIWSWCPVESGVPSPLLAQAQNFILTKN 166
Db 122 CTAGEAGTHNGKIGTCRCVAFNGSVRTCEIFAWCPVEVDTPFNPLLKEAENFIIFKN 181

QY 167 TVTFSKFNFSKNALETWPTVFKHCRYPQSPYCPVFRIGDLVAKAGTFEDLALGG 226
Db 182 SIRFPKFNFSKGNLENKTDLYLKHCRFHTNDPYCPIFRLDGVVEKAGQDFQDLAKGG 241

QY 227 SVGIRVHWDCLDTCGSCWPHYSFQ----LQES-----YNFRTHHWEOQGVARTLL 278
Db 242 VIGIINWDCDDDKAAECNPHYSFRRLDNKKEKSVSPGYNFRFAKYRDNNNGVEYRLL 301

QY 279 KLYGIRFDILVTGQAGKFGILPTAVTLGTGAALGVWTFECDLLLXYLDREAHFYWRKY 338
Db 302 KAYGIRFDVLNGKAGKFDIITINIGSLASLGVLGFLCDLILLYFLKKEHFYRDKKF 361

QY 339 EE 340
Db 362 EE 363

RESULT 4
US-10-128-558-144
; Sequence 144, Application US/10128558
; Publication No. US20040219521A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Zhiwei
; APPLICANT: Keng, Gezhi
; APPLICANT: Boyle, Bryan J
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/128,558
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: US 60/339,453
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 412
; SOFTWARE: PT_FL_genes Version 6.0
; SEQ ID NO 144
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-128-558-144

Query Match      37.1%; Score 825.5; DB 17; Length 422;
Best Local Similarity 42.7%; Pred. No. 1.7e-73;
Matches 166; Conservative 63; Mismatches 125; Indels 35; Gaps 7;

QY 20 WALLAKKGQERDLEPOFSIITKLKGSVSTQIKELGNRLMDVADVPKPOGENVFFLVN 79
Db 46 WVFVWEKGYQETD-SVVSSTVTKVGVAVNTSKLGFRIWDVADYVIPAQENSLSFVMTN 104

QY 80 FLVTPAQVQGRCPHPSPVLANCWVDEDCEBEGGTHSHGVKTCQCVVNGTHRTCEIWS 139
Db 105 VILTMNQTQGLCPPIPDATTV-CKSDASCTAGSAGTHSNGVSTGRCVAFNGSVKTCVAA 163

QY 140 WCPVESGV-VPSRPLLAQAQNFILTKNTVTFESKFNFSKNALETWPTVFKHCRYPQF 198
Db 164 WCPVEDDTHVPOPAFLKAAENFILLVKNNIWYKFNFSKENILPNITTTLYKSCIYDAKT 223

QY 199 SPYCPVFRIGDLVAKAGTFEDLALGGSVGIRVHWDCLDTCGSCWPHYSFQLE--- 255
Db 224 DPFCEPIFRLGKIVENAGHSQDMAVEGGIMGIQVWNCNLDRAASLCLPRYSFRRLTRD 283

QY 256 -----KSNFRTHHWEOQGVARTLLKLYGIRFDILVTGQAGKFGILPTAVTLGTGA 309
Db 284 VEHNVSGYGNFRFAKYRDLAGNEQRTLIKAYGIRFDIIVFGKAGKFDIIPITMINIGSL 343

QY 310 AWLGWVTFECDLLLXYLDREAHFYWRKYBEAKAPKATANSVWRELALASQARLAECLE 369
Db 344 ALLGMATVLCDDIIVLYCMKKLYREKKYKVVEDYEQVGPS-W----- 385

QY 370 SSAPAPTATAA---GSQTQTPGHPGPSDD 395
Db 386 --PPAGTGLSSLGCGSEGRGPMVGGPSAE 412

RESULT 5
US-09-833-082-2
; Sequence 2, Application US/09833082
; Patent No. US20020151480A1
; GENERAL INFORMATION:
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; FILE REFERENCE: CARDIOVASCULAR DISEASE USING 10218
; FILE REFERENCE: WHI-227
; CURRENT APPLICATION NUMBER: US/09/833,082
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-082-2

Query Match      36.7%; Score 816; DB 9; Length 388;
Best Local Similarity 47.0%; Pred. No. 1.4e-72;
Matches 155; Conservative 56; Mismatches 107; Indels 12; Gaps 4;

QY 20 WALLAKKGQERDLEPOFSIITKLKGSVSTQIKELGNRLMDVADVPKPOGENVFFLVN 79
Db 46 WVFVWEKGYQETD-SVVSSTVTKVGVAVNTSKLGFRIWDVADYVIPAQENSLSFVMTN 104

QY 80 FLVTPAQVQGRCPHPSPVLANCWVDEDCEBEGGTHSHGVKTCQCVVNGTHRTCEIWS 139
Db 105 VILTMNQTQGLCPPIPDATTV-CKSDASCTAGSAGTHSNGVSTGRCVAFNGSVKTCVAA 163
```

```

QY 140 WCPVSGV-VPSRPLLAQAQNTFLFKNTVTFKSNFKNLSNALETWDPYFKHCYEPQF 198
DB 164 WCPVEDDTHVQPAFLKAAENFTLLVKNINWPKFNFSKRNILPNTITTYLKSCYIDAKT 223
QY 199 SPYCPVERIGDLVAKAGTFFDLALLGSGVGIRVHWDCLDLDGSGCWPHYSFQLOE--- 255
DB 224 DFCPIFLGKIVENAGHSFQDMAVEGGIMGIQVNDONLDRASLCLPRYSFRFLDTRD 283
QY 256 -----KSNFRTATHWEPQVEARTLLKLYGIRFDILVTGQAGKFGLIPTAVTLGTGA 309
DB 284 VEHNVPGNFRFAKYRDLAGNEORTLTKAYGIRFDIIVFGKAGKFDIIPMTINIGSGL 343
QY 310 AWLGVTTFCDLLLVVDREAHFYWRKYE 339
DB 344 ALLGMATVLCDIIVLYCMKRLIYREKKYK 373

```

```

RESULT 6
US-10-455-552-2
; Sequence 2, Application US/10455552
; Publication No. US20040018533A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Gail Isabel
; APPLICANT: Langdown, Maria
; APPLICANT: Roth, Richard
; APPLICANT: Denissenko, Mikhail
; APPLICANT: Smylie, Kevin
; TITLE OF INVENTION: DIAGNOSING PREDISPOSITION TO FAT
; TITLE OF INVENTION: DEPOSITION AND THERAPEUTIC METHODS FOR REDUCING FAT
; TITLE OF INVENTION: DEPOSITION AND TREATMENT OF ASSOCIATED CONDITIONS
; FILE REFERENCE: 52459-20030.00
; CURRENT APPLICATION NUMBER: US/10/455,552
; PRIOR FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: US 60/386,012
; PRIOR FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-455-552-2

```

```

Query Match 36.7%; Score 816; DB 15; Length 388;
Best Local Similarity 47.0%; Pred. No. 1.4e-72;
Matches 155; Conservative 56; Mismatches 107; Indels 12; Gaps 4;

QY 20 WALLAKKGQERDLPEQPSIITKLKGVSVTQIKELGNRLWDVADVPKPPQGENVFLVTN 79
DB 46 WYFVWKEGQETD-SVSVSVTTKVGVAVTNTSKLGRINDVADYVIPAQEENSLFVMTN 104
QY 80 FLVTPAQVQGRCPHPSVPLANCWVDEDCPEGEGTHSHGVKTGCQVVFNGTHRTCEIWS 139
DB 105 VILTMNQTOGLCPEDATTV-CKSDASCTAGSAGTHSNGVSTGRCAVFNAGSVKTCVAA 163
QY 140 WCPVSGV-VPSRPLLAQAQNTFLFKNTVTFKSNFKNLSNALETWDPYFKHCYEPQF 198
DB 164 WCPVEDDTHVQPAFLKAAENFTLLVKNINWPKFNFSKRNILPNTITTYLKSCYIDAKT 223
QY 199 SPYCPVERIGDLVAKAGTFFDLALLGSGVGIRVHWDCLDLDGSGCWPHYSFQLOE--- 255
DB 284 VEHNVPGNFRFAKYRDLAGNEORTLTKAYGIRFDIIVFGKAGKFDIIPMTINIGSGL 343
QY 310 AWLGVTTFCDLLLVVDREAHFYWRKYE 339
DB 344 ALLGMATVLCDIIVLYCMKRLIYREKKYK 373

```

RESULT 7

```

US-10-386-414-17
; Sequence 17, Application US/10386414
; Publication No. US20040006016A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Robison, Keith E.
; APPLICANT: White, David
; APPLICANT: Williamson, Mark W.
; APPLICANT: Cook, William James
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: NOVEL 27875, 22025, 27420, 17906, 16319,
; TITLE OF INVENTION: 55092 AND 10218 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0210WNIM
; CURRENT APPLICATION NUMBER: US/10/386,414
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 09/426,282
; PRIOR FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 09/668,266
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/330,970
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: 09/724,599
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/860,193
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 10/283,023
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 60/335,044
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 10/010,943
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/254,037
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/833,082
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(388)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-386-414-17

```

```

Query Match 36.4%; Score 810; DB 15; Length 388;
Best Local Similarity 46.7%; Pred. No. 5.4e-72;
Matches 154; Conservative 56; Mismatches 108; Indels 12; Gaps 4;

QY 20 WALLAKKGQERDLPEQPSIITKLKGVSVTQIKELGNRLWDVADVPKPPQGENVFLVTN 79
DB 46 WYFVWKEGQETD-SVSVSVTTKVGVAVTNTSKLGRINDVADYVIPAQEENSLFVMTN 104
QY 80 FLVTPAQVQGRCPHPSVPLANCWVDEDCPEGEGTHSHGVKTGCQVVFNGTHRTCEIWS 139
DB 105 VILTMNQTOGLCPEDATTV-CKSDASCTAGSAGTHSNGVSTGRCAVFNAGSVKTCVAA 163
QY 140 WCPVSGV-VPSRPLLAQAQNTFLFKNTVTFKSNFKNLSNALETWDPYFKHCYEPQF 198
DB 164 WCPVEDDTHVQPAFLKAAENFTLLVKNINWPKFNFSKRNILPNTITTYLKSCYIDAKT 223
QY 199 SPYCPVERIGDLVAKAGTFFDLALLGSGVGIRVHWDCLDLDGSGCWPHYSFQLOE--- 255
DB 224 DFCPIFLGKIVENAGHSFQDMAVEGGIMGIQVNDONLDRASLCLPRYSFRFLDTRD 283
QY 256 -----KSNFRTATHWEPQVEARTLLKLYGIRFDILVTGQAGKFGLIPTAVTLGTGA 309
DB 284 VEHNVPGNFRFAKYRDLAGNEORTLTKAYGIRFDIIVFGKAGKFDIIPMTINIGSGL 343

```





QY 222 ALLGSGVIRVHWDCLDLDGSDGCPHYSTFQLOE-----KSYNFTATHWWEQPGVEA 274  
DB 244 AHKGGVIGVIINWDCDLDPASECNPKYSFRRLDPKHVPASSGYNRFFA-KYKINGTTT 302  
QY 275 RTLLKLVGIRDLVLTGQAGKFGILPTAVTLGTGAAMLGVVTFPCDLLLLVVDREAHFYW 334  
DB 303 RTLLKAYGIRDIVHVGQAGKFSUPTIINATALTSGVGSFLCDWILLTFMKNKQVYS 362  
QY 335 RTKYEEAKAPKATANSVWRELALASQARLAECLEERS--SAPAPTATAAGSQ-----TOT 386  
DB 363 HKKFDKVCPTSPHSGSNPVTLLARVLGQAPPBPGHRSDDQHPSPSGGQGGQGAECGPAPP 422  
QY 387 PGWPCPSS 394  
DB 423 PLRCPIS 430

RESULT 10  
US-10-051-874-123  
; Sequence 123, Application US/10051874  
; Publication No. US20040005557A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Coleman, Steven D  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Vernet, Corine AM  
; APPLICANT: Li, Li  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Guo, Xiaojia Sasha  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: MacDougall, John R  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Pena, Carol EA  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Zethusen, Bryan D  
; APPLICANT: Milliet, Isabelle  
; APPLICANT: Miller, Charles E  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Baumgartner, Jason C  
; APPLICANT: Herrman, John L  
; APPLICANT: Peyman, John A  
; APPLICANT: Gorman, Linda  
; APPLICANT: Mezes, Peter D  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Grosse, William M  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Rothenberg, Mark  
; APPLICANT: Stone, David J  
; APPLICANT: Burgess, Catherine E  
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF  
; FILE OF INVENTION: USING THE SAME  
; FILE REFERENCE: 21402-245  
; CURRENT APPLICATION NUMBER: US/10/051, 874  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: 60/268,595  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/325,306  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/262,587  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/272,409  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/262,454

QY 10 WGLL-DYKTEK-----WALLAKKGYQERDLEPOFSIITK 42  
DB 10 WSALWDYETPKVIVVNRRLGVLYRAVOLLLILLYFVWVYFVQKSYQSETPGESSIITK 69  
QY 43 LKGVSVTQIKELGNELWDVADPVKPPQGENVFFLVNLFVTPAQVQGRCPPEHPSVPLANC 102  
DB 70 VKGITSE-----HKVMDVEEVKPEGGSVFSIITRVEATHSQTQCTPESIRVHNATC 124  
QY 103 WYDECPGEGGTHSHGVKTQCV-VFNGHRTCTEIMWCPSVSVPSRLLAQAOHFT 161  
DB 125 LSDADCVAGELDMLGNLRTGSCVYQYQPSKTCVFCWCPVEDGASVQSLGWNAPNFT 184  
QY 162 LPIKNTVTFKFNFSKNALETWDTYFKHCRYEPQFSPYCPVFRIGDLVAKAGTPEDL 221  
DB 185 ILIKNSIHPKPHFKGNIAETD-GYLKRCCTFHASLDYCPFKLGFIVEKAGESFTEL 243  
QY 222 ALLGSGVIRVHWDCLDLDGSDGCPHYSTFQLOE-----KSYNFTATHWWEQPGVEA 274  
DB 244 AHKGGVIGVIINWDCDLDPASECNPKYSFRRLDPKHVPASSGYNRFFA-KYKINGTTT 302  
QY 275 RTLLKLVGIRDLVLTGQAGKFGILPTAVTLGTGAAMLGVVTFPCDLLLLVVDREAHFYW 334  
DB 303 RTLLKAYGIRDIVHVGQAGKFSUPTIINATALTSGVGSFLCDWILLTFMKNKQVYS 362  
QY 335 RTKYEEAKAPKATANSVWRELALASQARLAECLEERS--SAPAPTATAAGSQ-----TOT 386  
DB 363 HKKFDKVCPTSPHSGSNPVTLLARVLGQAPPBPGHRSDDQHPSPSGGQGGQGAECGPAPP 422  
QY 387 PGWPCPSS 394  
DB 423 PLRCPIS 430

Query Match 31.9%; Score 709; DB 15; Length 459;  
Best Local Similarity 36.9%; Pred. No. 8.9e-62;  
Matches 158; Conservative 66; Mismatches 154; Indels 50; Gaps 9;  
QY 10 WGLL-DYKTEK-----WALLAKKGYQERDLEPOFSIITK 42  
DB 10 WSALWDYETPKVIVVNRRLGVLYRAVOLLLILLYFVWVYFVQKSYQSETPGESSIITK 69  
QY 43 LKGVSVTQIKELGNELWDVADPVKPPQGENVFFLVNLFVTPAQVQGRCPPEHPSVPLANC 102  
DB 70 VKGITSE-----HKVMDVEEVKPEGGSVFSIITRVEATHSQTQCTPESIRVHNATC 124  
QY 103 WYDECPGEGGTHSHGVKTQCV-VFNGHRTCTEIMWCPSVSVPSRLLAQAOHFT 161  
DB 125 LSDADCVAGELDMLGNLRTGSCVYQYQPSKTCVFCWCPVEDGASVQSLGWNAPNFT 184  
QY 162 LPIKNTVTFKFNFSKNALETWDTYFKHCRYEPQFSPYCPVFRIGDLVAKAGTPEDL 221  
DB 185 ILIKNSIHPKPHFKGNIAETD-GYLKRCCTFHASLDYCPFKLGFIVEKAGESFTEL 243  
QY 222 ALLGSGVIRVHWDCLDLDGSDGCPHYSTFQLOE-----KSYNFTATHWWEQPGVEA 274  
DB 244 AHKGGVIGVIINWDCDLDPASECNPKYSFRRLDPKHVPASSGYNRFFA-KYKINGTTT 302  
QY 275 RTLLKLVGIRDLVLTGQAGKFGILPTAVTLGTGAAMLGVVTFPCDLLLLVVDREAHFYW 334  
DB 303 RTLLKAYGIRDIVHVGQAGKFSUPTIINATALTSGVGSFLCDWILLTFMKNKQVYS 362  
QY 335 RTKYEEAKAPKATANSVWRELALASQARLAECLEERS--SAPAPTATAAGSQ-----TOT 386  
DB 363 HKKFDKVCPTSPHSGSNPVTLLARVLGQAPPBPGHRSDDQHPSPSGGQGGQGAECGPAPP 422  
QY 387 PGWPCPSS 394  
DB 423 PLRCPIS 430

US-10-051-874-123  
; Sequence 2202, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04

```
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2202
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2202

Query Match      31.6%; Score 704.5; DB 16; Length 397;
Best Local Similarity 36.4%; Pred. No. 2.1e-61;
Matches 156; Conservative 59; Mismatches 114; Indels 101; Gaps 11;

QY 12 LLDYKTEK-----WALLAKGYQERDLPOPSIITKLG 45
DB 13 LFDYKTEKVIAKKNKGLLYRLLOASILAYLVVWFLIKGYQDVDTLSQSAVITKVG 72
QY 46 VSVTOIKELGNLMDVADFKPPQGENVFELVTNFLTPTAQVQCRCPEHPSVPLANCWD 105
DB 73 VAFNTNSDLQRIWDVADYVIPAQNE-----GIPDGACSKD 108
QY 106 EDCPEGEGGTHSHGVTKQCQVFNHGTCTEISWCPVBSGVVPSRPLLAQAQNFTEL 164
DB 109 SDCHAGEAVTAGVGTGRCLRENLAGTCEIFAWCPLETSRRPEEPFLKEADFTIFI 168
QY 165 KNTVTSKFNFSKNALETWDPYFKHCRYEPQFSPYCPVFRIGDLVAKAGGTEDL 224
DB 169 KHIRPKFNFS-NVMDVKDRSFLKSCFPGK-NHYCFIRLGSVIRWAGSDFOIALE 226
QY 225 GGSVGRVHWDCDLDTGDSGCWPHYSF-QLOEK-----SYNFRATATHWQPGVEART 276
DB 227 GGIVIGINIEWNCDLRAASECHPHYSFSLDNKLSKSVSSGYNFRFARYRDAAGVEPRT 286
QY 277 LLKLYGIRDLVTGQAGKFLIPTAVTLTGTAAMLGVVTFECDLLLLVVDREAHFYWT 336
DB 287 LMKAYGIRDVHVGQAGKFSLIPTIINLATALTSVGVSFLCDWILLTFMKNKYVS 372
QY 337 KYEAKAPKATANSVWRELASOARLAECRLRSAPA-----PTATAAGSQ 383
DB 325 KYEEVRGLDSSQEADE--ASGLGLSEQL--TSGPGLGLMPEQQLQEPPEAKRGSS 379
QY 384 TQTPGQWPCP 392
DB 380 QKXGNSVCP 388

RESULT 12
US-09-764-875-905
; Sequence 905, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 905
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-905

Query Match      31.2%; Score 694; DB 11; Length 402;
Best Local Similarity 38.8%; Pred. No. 2.4e-60;
Matches 142; Conservative 63; Mismatches 119; Indels 42; Gaps 7;

QY 10 WCLL-DYKTEK-----WALLAKGYQERDLPOPSIITK 42
DB 20 WSAWDYETPKVIVVVRNRLGLYRAVQLILLYFVWVYVQKSYQSSRGPSIITK 79
QY 43 LKGVSVTQIKELGNLMDVADFKPPQGENVFELVTNFLTPTAQVQCRCPEHPSVPLANC 102

; RESULT 13
US-10-051-874-124
; Sequence 124, Application US/10051874
; Publication No. US2004000557A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Coleman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zernhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgarther, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Payman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
```

APPLICANT: Edinger, Shlomit R  
APPLICANT: MacDougall, John R  
APPLICANT: Malyankar, Uriel M  
APPLICANT: Patturajan, Meera  
APPLICANT: Shimkets, Richard A  
APPLICANT: Pena, Carol EA  
APPLICANT: Tchernev, Velizar T  
APPLICANT: Zernusen, Bryan D  
APPLICANT: Millet, Isabelle  
APPLICANT: Miller, Charles E  
APPLICANT: Lepley, Denise M  
APPLICANT: Smithson, Glenda  
APPLICANT: Baumgartner, Jason C  
APPLICANT: Herrman, John L  
APPLICANT: Feyman, John A  
APPLICANT: Gorman, Linda  
APPLICANT: Mezes, Peter D  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Taupier Jr, Raymond J  
APPLICANT: Gerlach, Valerie  
APPLICANT: Grosse, William M  
APPLICANT: Liu, Xiaohong  
APPLICANT: Ellerman, Karen  
APPLICANT: Rothenberg, Mark  
APPLICANT: Stone, David J  
APPLICANT: Burgess, Catherine E  
APPLICANT: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF  
TITLE OF INVENTION: USING THE SAME  
FILE REFERENCE: 21402-245  
CURRENT APPLICATION NUMBER: US/10/051,874  
CURRENT FILING DATE: 2002-09-25  
PRIOR APPLICATION NUMBER: 60/268,595  
PRIOR FILING DATE: 2001-02-14  
PRIOR APPLICATION NUMBER: 60/325,306  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: 60/262,587  
PRIOR FILING DATE: 2001-01-18  
PRIOR APPLICATION NUMBER: 60/272,409  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 60/262,454  
PRIOR FILING DATE: 2001-01-18  
PRIOR APPLICATION NUMBER: 60/276,777  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/291,672  
PRIOR FILING DATE: 2001-05-17  
PRIOR APPLICATION NUMBER: 60/330,336  
PRIOR FILING DATE: 2001-10-18  
PRIOR APPLICATION NUMBER: 60/265,530  
PRIOR FILING DATE: 2001-01-31  
PRIOR APPLICATION NUMBER: 60/261,376  
PRIOR FILING DATE: 2001-01-16  
NUMBER OF SEQ ID NOS: 269  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 120  
LENGTH: 497  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-051-874-120

Query Match 31.1%; Score 693; DB 15; Length 497;  
Best Local Similarity 35.0%; Pred. No. 4e-60;  
Matches 159; Conservative 66; Mismatches 153; Indels 76; Gaps 10;  
Qy 10 WGLL-DYKTEK-----WALLAKGYQERDLEPQFSIITK 42  
Db 22 WSALWDYETPKVIWVRNRRLGLVYRAVOLLLILYFWYVIVQKSYQESGTPSSIIITK 81  
Qy 43 LKGVSVTQIKELGNRLWDVADFKPQGNVFLVFNFLVTPAQVQGRCPHFSPVLANC 102  
Db 82 VKGIITSE-----HKVWDVEEVKPEGGSVFSIIITRVEATHSQTGTCTPESIRVHNATC 136  
Qy 103 WYDEDCPEGGGTHSHGVTGQCV-VFNGTHRTCEIWSWCPVSGVVPRLAQONFT 161

PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: 60/262,587  
PRIOR FILING DATE: 2001-01-18  
PRIOR APPLICATION NUMBER: 60/272,409  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 60/262,454  
PRIOR FILING DATE: 2001-01-18  
PRIOR APPLICATION NUMBER: 60/276,777  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/291,672  
PRIOR FILING DATE: 2001-05-17  
PRIOR APPLICATION NUMBER: 60/330,336  
PRIOR FILING DATE: 2001-10-18  
PRIOR APPLICATION NUMBER: 60/265,530  
PRIOR FILING DATE: 2001-01-31  
PRIOR APPLICATION NUMBER: 60/261,376  
PRIOR FILING DATE: 2001-01-16  
NUMBER OF SEQ ID NOS: 269  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 124  
LENGTH: 404  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-051-874-124

Query Match 31.2%; Score 694; DB 15; Length 404;  
Best Local Similarity 38.8%; Pred. No. 2.4e-60;  
Matches 142; Conservative 63; Mismatches 119; Indels 42; Gaps 7;  
Qy 10 WGLL-DYKTEK-----WALLAKGYQERDLEPQFSIITK 42  
Db 22 WSALWDYETPKVIWVRNRRLGLVYRAVOLLLILYFWYVIVQKSYQESGTPSSIIITK 81  
Qy 43 LKGVSVTQIKELGNRLWDVADFKPQGNVFLVFNFLVTPAQVQGRCPHFSPVLANC 102  
Db 82 VKGIITSE-----HKVWDVEEVKPEGGSVFSIIITRVEATHSQTGTCTPESIRVHNATC 136  
Qy 103 WYDEDCPEGGGTHSHGVTGQCV-VFNGTHRTCEIWSWCPVSGVVPRLAQONFT 161  
Db 137 LSDADCVAGELMGLNGURTGSCVYQYQPSKTCVFVGCWVEDGASVQLGTWAPNFT 196  
Qy 162 LPIKNTVTFKSNFKNALETWDTYFKHCRIYEPQFSPYCFVFRIGDLVAKAGTFEDL 221  
Db 197 ILIKNSIHYKPHFKSGNIADRTD-GYLKRCCTFHEASLDYCPFKLGFIVEKAGESFTL 255  
Qy 222 ALLGSVGRVFWDCDLDLTDGSGCWPHYSFQLE-----KSNFRATTHWQPGVEA 274  
Db 256 AHKGGVIGVIINWDCDLDPASECNPKYSFRLLDPKHPVASSGYNFRFA-KYYKINGTTT 314  
Qy 275 RTLLKLYGIRFOLVITGQAKGFLIPTAVTLGTGAALGWVTFPCDLLLLLYVDREAHFYW 334  
Db 315 RLIIKAYGIRIDIVHGQAGKFSLIPTIINLATLTSVGVGSLCDWILLTFMKNKYVS 374  
Qy 335 RTKYE 340  
Db 375 HKKFDK 380

RESULT 14  
US-10-051-874-120  
Sequence 120, Application US/10051874  
Publication No. US2004000557A1  
GENERAL INFORMATION:  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Alsebrook II, John P  
APPLICANT: Colman, Steven D  
APPLICANT: Spytex, Kimberly A  
APPLICANT: Boldog, Ferenc  
APPLICANT: Vernet, Corine AM  
APPLICANT: Li, Li  
APPLICANT: Shenoy, Suresh G  
APPLICANT: Casman, Stacie J  
APPLICANT: Guo, Xiaojia Sasha

Search completed: November 21, 2004, 15:15:34  
Job time : 146 secs

```
Db 137 LSDADCVAGELDMGLNGLRTGRCVPYVQGPSKTCVEFGWCPVEDGASVSQFLGTMAPNET 196
QY 162 LFIKNVTWTSKFNFSKNALETWDPYFKHCRVEPOFSPYCPVFRIGDLVAKAGTFEDL 221
Db 197 ILKNSIHVPKPHFSKGNADRTD-GYLKRCFTPEASDLVCPFKLGFIVEKAGESFTEL 255
QY 222 ALLGSGVIRVHWDCCDLDTGSCWPHYSFQLOE-----KSYNFTATHWMEQPGVEA 274
Db 256 AHKGGVIGVIINWDCDLDPASECNPXYFRRLDPKHVEASSGYNERFA-KYIKINGTTT 314
QY 275 RTLLKLYGIRFDILVTGQAGKFGLIPTAVTLGTGAANLGV----- 315
Db 315 RTLIKAYGIRIDVIVHGOAGKFSLIPTIINLATITSVGVWRNPLWGPSCGGSTRPLHT 374
QY 316 -----TFECDLLLYVDREAHFYWRTKYVEAKAPKATANSWRELALASQARLAECUR 368
Db 375 GLCWPGSFLCDWILLTFMNKKNVSHKFKDVCFTSPHSGSNPVTLARVLGOAPPEPGH 434
QY 369 RS--SAPAPTATAAGSQ-----TQTPGWFPCPS 394
Db 435 RSEDQHPSPSPSGQGQGAECGFAFPFLRCPIS 468

RESULT 15
US-10-455-552-3
; Sequence 3, Application US/10455552
; Publication No. US20040018533A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Gail Isabel
; APPLICANT: Langdown, Maria
; APPLICANT: Roth, Richard
; APPLICANT: Denissenko, Mikhail
; APPLICANT: Smylie, Kevin
; TITLE OF INVENTION: DIAGNOSING PREDISPOSITION TO FAT
; TITLE OF INVENTION: DEPOSITION AND THERAPEUTIC METHODS FOR REDUCING FAT
; TITLE OF INVENTION: DEPOSITION AND TREATMENT OF ASSOCIATED CONDITIONS
; FILE REFERENCE: 52459-20030.00
; CURRENT APPLICATION NUMBER: US/10/455,552
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: US 60/386,012
; PRIOR FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-455-552-3

Query Match 30.3%; Score 673.5; DB 15; Length 287;
Best Local Similarity 46.2%; Pred. No. 1.7e-58;
Matches 126; Conservative 47; Mismatches 89; Indels 11; Gaps 3;

QY 77 VTNFLTVAQVQGRCPHPSVPLANCWVDEDCPEGEGGTHSHGVKTCGVVFNHTRTCE 136
Db 1 MTNVLITWNTQGLQPEIPDATTV-CKSDASCTAGSAGTHSNGVSTGRCAVFNHGVKICE 59
QY 137 IWSWCPVBSGV-VPSRPLLAQONFTLFIKNTVTSKFNFSKNALETWDPYFKHCKYE 195
Db 60 VAAWCEVEDDTHVQPAFLKAAENFTLLVXNINWPKFNFSKRNILPNITTYLKSCTYD 119
QY 196 POFSPYCPVFRIGDLVAKAGTFEDLALLGSGVIRVHWDCCDLDTGDSGWPHYSFQLOE 255
Db 120 AKTDFPCIFRGLKIVENAGHSFQMAVEGGIMGQVNDNDRRAASLCILPRYSFRRLD 179
QY 256 -----KSYNFTATHWMEQPGVEARTLLKLYGIRFDILVTGQAGKFGLIPTAVTLG 306
Db 180 TRDVEHNVSPGYNFRFAKYRDLAGNEQRTLIKAYGIRFDIIVFGKAGKFDIITMINIG 239
QY 307 TGAANLGVVTFECDLLLYVDREAHFYWRTKYE 339
Db 240 SCLALLGNATVLCDIIVLYCMKKRLIYREKKYK 272
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:54:27 ; Search time 40 Seconds  
(without alignments)  
671.470 Million cell updates/sec

Title: US-09-820-095B-2

Perfect score: 2226

Sequence: 1 MGSPGATTGGLLDYKTEK.....TFGWPCPSDTHLPHSGSL 405

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/aa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/1/aa/5B-COMB.pep:\*  
3: /cgn2\_6/prodata/1/aa/5A-COMB.pep:\*  
4: /cgn2\_6/prodata/1/aa/5B-COMB.pep:\*  
5: /cgn2\_6/prodata/1/aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/prodata/1/aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2203	99.0	431	3	US-09-381-681-3
2	2203	99.0	441	3	US-09-191-136-31
3	1854	69.8	379	3	US-09-191-136-32
4	816	36.7	388	2	US-08-742-621-1
5	813	36.5	388	3	US-09-191-608-22
6	794	35.7	388	2	US-08-750-134A-7
7	794	35.7	388	3	US-09-363-745-7
8	793	35.6	388	3	US-09-191-608-23
9	744	33.4	399	2	US-08-742-621-3
10	744	33.4	399	3	US-08-750-134A-11
11	744	33.4	399	3	US-09-363-745-11
12	739	33.2	399	2	US-08-742-621-4
13	739	33.2	399	2	US-08-750-134A-5
14	739	33.2	399	3	US-09-363-745-5
15	720	32.3	472	2	US-08-742-621-5
16	720	32.3	472	3	US-08-842-079-15
17	720	32.3	472	3	US-09-638-857-15
18	709	31.9	471	3	US-09-191-608-17
19	694	31.2	404	3	US-09-191-608-18
20	693	31.1	497	3	US-09-191-608-20
21	679.5	30.5	397	2	US-08-750-134A-9
22	679.5	30.5	397	3	US-09-363-745-9
23	679.5	30.5	397	3	US-09-191-136-17
24	673.5	30.3	397	3	US-09-191-136-16
25	631	28.3	447	3	US-09-191-608-19
26	608.5	27.3	595	3	US-08-842-079-18
27	608.5	27.3	595	3	US-08-842-079-20

Sequence 18, Appl  
Sequence 20, Appl  
Sequence 6, Appl  
Sequence 17, Appl  
Sequence 6, Appl  
Sequence 17, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 7099, Ap  
Sequence 43, Appl  
Sequence 69, Appl  
Sequence 11, Appl  
Sequence 12, Appl  
Sequence 11711, A

28 608.5 27.3 595 4 US-09-638-857-18  
29 608.5 27.3 595 4 US-09-638-857-20  
30 602.5 27.1 595 3 US-08-842-079-6  
31 602.5 27.1 595 3 US-08-842-079-17  
32 602.5 27.1 595 4 US-09-638-857-6  
33 602.5 27.1 595 4 US-09-638-857-17  
34 349 15.7 211 1 US-07-915-934-4  
35 349 15.7 211 1 US-08-325-743-4  
36 95.5 4.3 851 4 US-09-543-681A-7277  
37 95 4.3 1041 3 US-08-898-978-2  
38 95 4.3 1041 3 US-09-372-858-2  
39 89.5 4.0 389 4 US-09-328-352-7099  
40 87 3.9 451 4 US-09-578-063-43  
41 87 3.9 470 4 US-09-578-063-69  
42 86.5 3.9 352 2 US-08-483-926A-11  
43 86 3.9 631 3 US-09-345-468-12  
44 86 3.9 631 3 US-09-414-453A-12  
45 85.5 3.8 827 4 US-09-489-039A-11711

#### ALIGNMENTS

RESULT 1  
US-09-381-681-3  
; Sequence 3, Application US/09381681  
; Patent No. 6255472  
; GENERAL INFORMATION:  
; APPLICANT: TAKINO, Takashi  
; APPLICANT: NAKAMURA, Yusuke  
; TITLE OF INVENTION: HUMAN GENES  
; FILE REFERENCE: Q55876  
; CURRENT APPLICATION NUMBER: US/09/381.681  
; CURRENT FILING DATE: 2000-01-10  
; EARLIER APPLICATION NUMBER: JPA 9-093044  
; EARLIER FILING DATE: 1997-03-26  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-381-681-3

Query Match 99.0%; Score 2203; DB 3; Length 431;  
Best Local Similarity 94.0%; Pred. No. 2.2e-235;  
Matches 405; Conservative 0; Mismatches 0; Indels 26; Gaps 1;  
Qy 1 MGSPGATTGGLLDYKTEK-----WALLAKGYQERDLE 34  
Db 1 MGSPGATTGGLLDYKTEKYNTRNRVVGALQLLQFGIVVYVVGWALLAKGYQERDLE 60  
Qy 35 PQFSIIITKLKGVSVTQIKELGNRLWDVADFVKPPQGENVFFLVNFTLTPAQVQGRCPHEH 94  
Db 61 PQFSIIITKLKGVSVTQIKELGNRLWDVADFVKPPQGENVFFLVNFTLTPAQVQGRCPHEH 120  
Qy 95 PSVPLANCWDEDCPGEGETSHGKVTGCQVFNCTHETCEIWSNCPVSGVVPSPRL 154  
Db 121 PSVPLANCWDEDCPGEGETSHGKVTGCQVFNCTHETCEIWSNCPVSGVVPSPRL 180  
Qy 155 AQQNFTLFKNVTFPSKFNFSKNALETWDPYFKHCRVEPOFSPYCPVFRIGDLVAKA 214  
Db 181 AQQNFTLFKNVTFPSKFNFSKNALETWDPYFKHCRVEPOFSPYCPVFRIGDLVAKA 240  
Qy 215 GGTFFEDIALGSGVGRVHWDCDLDTGDSGCPHYSFQLQKSYNFRTHHWEQGVGA 274  
Db 241 GGTFFEDIALGSGVGRVHWDCDLDTGDSGCPHYSFQLQKSYNFRTHHWEQGVGA 300  
Qy 275 RTLLKLYGIRFDILVTGQAKGFLIPTAVTLGTGAALGVWTFVFCOLLILLYVDREAHFYW 334  
Db 301 RTLLKLYGIRFDILVTGQAKGFLIPTAVTLGTGAALGVWTFVFCOLLILLYVDREAHFYW 360  
Qy 335 RTKYBEAKPAPKATANSVWRELALASQARLAELRSSAPAPTATAAGSQTPTGWPCCPS 394

Db 431 DTHLPHSGSL 441

RESULT 3

US-09-191-136-32

; Sequence 32, Application US/09191136B

; Patent No. 6214581

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Lynch, Kevin J.

; APPLICANT: Burgard, Edward C.

; APPLICANT: Van Biesen, T.

; TITLE OF INVENTION: Nucleic Acids Encoding A Functional

; TITLE OF INVENTION: Human Purinoreceptor P2X3 and P2X6 And Methods Of Production

; TITLE OF INVENTION: And Use Thereof

; FILE REFERENCE: 6293.US.P1

; CURRENT APPLICATION NUMBER: US/09/191,136B

; EARLIER FILING DATE: 1998-11-13

; EARLIER APPLICATION NUMBER: US 09/008,526

; EARLIER FILING DATE: 1998-01-16

; EARLIER APPLICATION NUMBER: US 09/008,185

; EARLIER FILING DATE: 1998-01-16

; EARLIER APPLICATION NUMBER: US 60/071,298

; EARLIER FILING DATE: 1998-01-16

; EARLIER APPLICATION NUMBER: US 60/071,669

; EARLIER FILING DATE: 1998-01-16

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 32

; LENGTH: 379

; TYPE: PRT

; ORGANISM: Rattus rattus

US-09-191-136-32

Query Match 69.8%; Score 1554; DB 3; Length 379;

Best Local Similarity 75.7%; Pred. No. 1.8e-163;

Matches 281; Conservative 26; Mismatches 38; Indels 26; Gaps 1;

QY 6 ATTGGLLDYKTEK-----WALLAKGYQERDLEPQPSI 39

Db 8 ALVSWGFLDYKTEKYVMTRNCWVGISQRLQLGVVVVIGWALLAKGYQEWMDPQISV 67

QY 40 ITHKGVSVTQIKELGNRLMDVADFKVPPQGENVFLVNTFLVTPAQQGRCPEHPSVPL 99

Db 68 ITHKGVSVTQIKELGNRLMDVADFKVPPQGENVFLVNTFLVTPAQQGRCPEHPSVPL 127

QY 100 ANCWVEDCPGEGGTHSHGVKTGQCVFNGTHRTCEIWSWCPVSGVVPSPPLAQAKN 159

Db 128 ANCWVEDCPGEGGTHSHGVKTGQCVFNGTHRTCEIWSWCPVSGVVPSPPLAQAKN 187

QY 160 FTLFIKNTVTFKFNFSKNALETWDPPTYFKHCRVEPQSPYCPVFRIGDLVAKAGGTFE 219

Db 188 FTLFIKNTVTFKFNFSKNALETWDPPTYFKHCRVEPQSPYCPVFRIGDLVAKAGGTFE 247

QY 220 DLALLGGSGIRVHWDCLDTGDSGCPHYSFQLOEKSYNFRATATHWEPQGVARTLLK 279

Db 248 DLALLGGAGINHWDCNLDTKGDCSPQSYFQLOEKGYNFRATANYWMAASGVESRSLK 307

QY 280 LYGIREDILVTGQAGKFGILPTAVTLGTGAAMLVGVVTFPCDILLVYDREAHFYWTKYE 339

Db 308 LYGIREDILVTGQAGKFGILPTAVTLGTGAAMLVGVVTFPCDILLVYDREAHFYWTKYE 367

QY 340 EAKAPKATANS 350

Db 368 EAKAPKATANS 378

RESULT 4

US-08-742-621-1

; Sequence 1, Application US/08742621

; Patent No. 5856129

; GENERAL INFORMATION:

; APPLICANT: HILLMAN, JENNIFER L.



RESULT 5  
US-09-191-608-22  
; Sequence 22, Application US/09191608  
; Patent No. 6242216

RESULT 6  
US-08-750-134A-7  
Sequence 7, Application US/08750134A  
Patent No. 5985603  
GENERAL INFORMATION:  
APPLICANT: VALERA, SOLEDAD  
APPLICANT: BUELL, GARY  
TITLE OF INVENTION: P2x RECEPTORS (PURENOCEPTOR FAMILY)  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,134A



QY 20 WALLAKGYOERDLBPQSIITKLKGVSVTOIKELGNRLWDVADFKVPPQGENVFLVTN 79  
Db 46 WVFWEKGYOERDLBPQSIITKLKGVSVTOIKELGNRLWDVADFKVPPQGENVFLVTN 104  
QY 80 FLVTPAQVQGRCPHPSVPLANCWVDEDCPEGEGGTHSHGVKTCQCVVFNHGTHTCEIWS 139  
Db 105 MIVTVNQTSTCPEIPD-KTSLCNSDADCTPGSVDTSHSGVATGCVFNFESVKTCEVAA 163  
QY 140 WCPVPSGV-VPSRPLLAQONFTLFIKNVTFTSKFNPSKNALETWDTPTFKHCRYPQF 198  
Db 164 WCPVENDGVPTPAFLKAAENFTLLVKNINWPKFNPSKRNLTPLNITTSYLKSCINYAQT 223  
QY 199 SPYCPVFRIGDLVAKAGGTFFDLALLGGSGVIRVHWDCLDGTGSGCWPHYSFQJQE--- 255  
Db 224 DFPCIFRLGTIVEDAGHSFQEWAVEGGIMGQIKWDCNLDRAASLCPLPRYSFRLDTRD 283  
QY 256 -----KSNFTATHWEPQGVARTLLKLYGIRFDILVTQAGKFGILPTAVTLTGGA 309  
Db 284 LEHNVSPGYNFRFAKYRDLAGKQRTLTAKYGIRFDIIVFGKAGKFDIIPMTINVGSL 343  
QY 310 AMLGVVTFPCDLLLLLVYDREAHFYWRTKYE 339  
Db 344 ALLGVATVLCVIVLYCMKKYKYYDKYK 373

RESULT 9  
US-08-742-621-3  
; Sequence 3, Application US/08742621  
; Patent No. 5856129  
; GENERAL INFORMATION:  
; APPLICANT: HILLMAN, JENNIFER L.  
; APPLICANT: COLEMAN, ROGER  
; TITLE OF INVENTION: NOVEL HUMAN PURINOCEPTOR  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/742,621  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PP-0147 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 399 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 166438  
US-08-742-621-3

Query Match 33.4%; Score 744; DB 2; Length 399;

Best Local Similarity 44.0%; Pred. No. 1.5e-73;  
Matches 159; Conservative 56; Mismatches 128; Indels 18; Gaps 9;  
QY 20 WALLAKGYOERDLBPQSIITKLKGVSVTOIKELGNRLWDVADFKVPPQGENVFLVTN 79  
Db 47 WVFWEKGYOERDLBPQSIITKLKGVSVTOIKELGNRLWDVADFKVPPQGENVFLVTN 105  
QY 80 FLVTPAQVQGRCPHPSVPLANCWVDEDCPEGEGGTHSHGVKTCQCVVFNHGTHTCEIWS 139  
Db 106 FIVTPKTOGCAEHPGGI--CKEDSGCTPGKAKRKAQGIKTKCVAFNDTVKTCFIFG 163  
QY 140 WCPVPSGV-VPSRPLLAQONFTLFIKNVTFTSKFNPSKNALETWDTPTFKHCRYPQF 198  
Db 164 WCPVEVDDDDIPRALLREAENFTLFIKNISISFFPKVNRNRENVBEVNAAMHMKTCLEHKL 223  
QY 199 SPYCPVFRIGDLVAKAGGTFFDLALLGGSGVIRVHWDCLDGTGSGCWPHYSFQ--LQEK 256  
Db 224 HPLCPVQLGVVQESQNFSTLAEXGVVGIIDWHICDLHVRHCRPIYEFHGLYEK 283  
QY 257 S-----YNFRATHWEPQGVARTLLKLYGIRFDILVTQAGKFGILPTAVTLTGGA 312  
Db 284 NLSGFGNFRFAHFVEN-GTNYRHLFKVFGIRFDILVDGKAGKFDIIPMTTIGSGIGIF 342  
QY 313 GVVTFFCDLLLLLVYDREAHFY--WRTKYEEAKAPKATANSVWRELALASQA-ELAECLRR 369  
Db 343 GVATVLCDDLLHLPKRHYKOKKFKYAEKMGAGAE-----RDLAATSTSLGQENMRT 398  
QY 370 S 370  
Db 399 S 399

RESULT 10  
US-08-750-134A-11  
; Sequence 11, Application US/08750134A  
; Patent No. 5985603  
; GENERAL INFORMATION:  
; APPLICANT: VALERA, SOLEDAD  
; APPLICANT: BUELL, GARY  
; TITLE OF INVENTION: P2X RECEPTORS (PURINOCEPTOR FAMILY)  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHIVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/750,134A  
; FILING DATE: 22-JAN-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAWFORD, ARTHUR C.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 1430-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4006  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 399 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-750-134A-11

Query Match 33.4%; Score 744; DB 2; Length 399;

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

```

US-09-363-745-11

Query Match      33.4%; Score 744; DB 3; Length 399;
Best Local Similarity 44.0%; Pred. No. 1.5e-73;
Matches 159; Conservative 56; Mismatches 128; Indels 18; Gaps 9

QY 20 WALLAKCYQERDLEPQFSIITKLKGVSVTQIKELGNRLWDVADFVKPQQGENVFLVTN 79
DB 47 WFLYKGYQTSS-GLISSVSUKLGLAVTQLPGLGQVMDVADYVFPQAGDSFVVMTN 105
QY 80 FLVTPAQVQGRCPHPSPVPLANCWVDEDCPEGGTHSHGVKTCQCWFNGTHRTCEIWS 139
DB 106 FIVTPKQTQYCAEHPEGGI--CKEDSGCTPGRAKKAQGI RTGKCVAFNDVTKTCIEFG 163
QY 140 WCPVZ-SGVVSRPLLQAQNFLLFKNTVTFKFNFSKNALETWDPTYFKHCRYEPQF 198
DB 164 WCPVEVDDIIPRALREAE NFLLFKNSISFPRFVKNRNLVEEVNAAHMTCTLFHKT 223
QY 199 SPYCVFVRIQVLVAKAGGTEDLALLGSGVIRVHWDCCDLDTGDSGWPHYSFQ--LQEK 256
DB 224 HPLCFVFLQGVVQESGQNFSTLAEKGGVVGIIIDHCDLDWHVHCRPIYFTHGLYEEK 283
QY 257 S-----YNPRTATHWBPQGVAEARTLLKLYGIRFDILVTGQAGHFGLLPTAVTLGTGAWL 312
DB 284 NLSQGFNFRFAHFVEN-GTNYRHLFKVFGIRFDILVDGKAGFDIIPMTTITGSGIGIF 342
QY 313 GVVTFFCDLLLLLYDRAHFY--WRTYVEAKAPKATANSVMRELALASOA--RLAECRLR 369
DB 343 GVATVLCDLLLHLIPRHHYKQKKFYAEDMGPGAEE----RDLAATSTSLGLQENMRT 398
QY 370 S 370
DB 399 S 399

```

```

1 RESULT 12
2 US-08-742-621-4
3 ; Sequence 4, Application US/08742621
4 ; Patent No. 5856129
5 ; GENERAL INFORMATION:
6 ; APPLICANT: HILLMAN, JENNIFER L.
7 ; APPLICANT: COLEMAN, ROGER
8 ; TITLE OF INVENTION: NOVEL HUMAN PURINOCEPTOR
9 ; NUMBER OF SEQUENCES: 5
10 ; CORRESPONDENCE ADDRESS:
11 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
12 ; STREET: 3174 Porter Drive
13 ; CITY: Palo Alto
14 ; STATE: CA
15 ; COUNTRY: US
16 ; ZIP: 94304
17 ; COMPUTER READABLE FORM:
18 ; MEDIUM TYPE: Diskette
19 ; COMPUTER: IBM Compatible
20 ; OPERATING SYSTEM: DOS
21 ; SOFTWARE: FastSeq Version 1.5
22 ; CURRENT APPLICATION DATA:
23 ; APPLICATION NUMBER: US/08/742,621
24 ; FILING DATE: Filed Herewith
25 ; PRIOR APPLICATION DATA:
26 ; APPLICATION NUMBER:
27 ; FILING DATE:
28 ; ATTORNEY/AGENT INFORMATION:
29 ; NAME: Billings, Lucy J.
30 ; REGISTRATION NUMBER: 36,749
31 ; REFERENCE/DOCKET NUMBER: PF-0147 US
32 ; TELECOMMUNICATION INFORMATION:
33 ; TELEPHONE: 415-855-0555
34 ; TELEFAX: 415-845-4166
35 ; INFORMATION FOR SEQ ID NO: 4:
36 ; SEQUENCE CHARACTERISTICS:
37 ; LENGTH: 399 amino acids
38 ; TYPE: amino acid

```



```

; MOLECULE TYPE: protein
; US-09-363-745-5
Query Match 33.2%; Score 739; DB 3; Length 399;
Best Local Similarity 43.8%; Pred. No. 5,4e-73;
Matches 153; Conservative 60; Mismatches 114; Indels 22; Gaps 8;
QY 20 WALLAKKGVQ-BRDLEPFSIITKLKGVSVTOIKELGNRLWDVADFKVPPQGENVFLVLT 78
DB 47 WYFVVEKGYQTSDDSISSVSU--KLGLAVTQLOGLPQVMDVADYVFPAPGDSFVVMVT 104
QY 79 NFLVTPAQVQGRCPHPSVPLANCWDEDCPEGEGTSHGKVTGOCV-VFNGTHTCTBIW 138
DB 105 NFIVTPQQTQGHCAENPEGGI--CQDSDGCTPGKAERKAQGIKRGNCVFPNGTCKEIP 162
QY 139 SWCPVE-SGVVPSRPLLAQAQNTFLFIKNTVTFKFNFSKNALETWDPDYFKHCRYEPQ 197
DB 163 GWCPEVEDDKIPSPALLREAENFTLFIKNSISFPFRKVNRENLEVEVNGTVMKCLYHKI 222
QY 198 FSPYCPVFRIGDLVAKAGTGFEDLALGSGVGIRVHWCDDLDTGDSGCWPHYSFQ--LQE 255
DB 223 QHPLCPVFNGLGVVRESQDFSLAEKGVVGIIIDWKCDLDWVRHCKPIYQFHLYGE 282
QY 256 KS-----YNFRTATHWVEQPGVEARTLLKLYGIRFDILVTGQAGKFGLIPTAVTLGTGA 311
DB 283 KNLSPGFNFRFAH-FVQNGTNRHLFKVFGIHFDILVDGKAGKFDIIPNTMTTIGSGIGI 341
QY 312 LGVTFPFCOLLILLYVDREAHFYWRKYBEAKPKATANSVWRELALASQARLAECLERSS 371
DB 342 FGVAIVLCDELLHLPKRHYKQKFKYAEQMGEGEHDPVATSTL 390

RESULT 15
US-08-742-621-5
Sequence 5, Application US/08742621
Patent No. 5856129
GENERAL INFORMATION:
APPLICANT: HILLMAN, JENNIFER L.
APPLICANT: COLEMAN, ROGER
TITLE OF INVENTION: NOVEL HUMAN PURINOCEPTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,621
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0147 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Gensbank
; CLONE: 558831
; US-08-742-621-5
Query Match 32.3%; Score 720; DB 2; Length 472;
Best Local Similarity 37.4%; Pred. No. 9e-71;
Matches 144; Conservative 71; Mismatches 144; Indels 26; Gaps 5;
QY 20 WALLAKKGVQERDLBPQFSIITKLKGVSVTOIKELGNRLWDVADFKVPPQGENVFLVLT 79
DB 47 YFVIVQSYQDSGTGPSSIIITKVKITMSE-----DKVMDVEEYKPEGSGSVSIITR 101
QY 80 FLVTPAQVQGRCPHPSVPLANCWDEDCPEGEGTSHGKVTGOCV-VFNGTHTCTBIW 138
DB 102 IEVTPSQTLTGPESVRVHSSCHSDDDCIAGQLDMQNGIRTGHCVPYHGDSTKCEVS 161
QY 139 SWCPVESGVVPSRPLLAQAQNTFLFIKNTVTFKFNFSKNALETWDPDYFKHCRYEPQ 198
DB 162 AWCPEVEDGTSNHLFLGKMAPNFTILIKNSIHYPKFKFSKGN-IASQKSDYLKXCTFDQDS 220
QY 199 SPYCPVFRIGDLVAKAGTGFEDLALGSGVGIRVHWCDDLDTGDSGCWPHYSFQ--LQE 255
DB 221 DYCPIFRIGFIVEKAGENFTLAHKGVGIVGIINWNCDDLDSECECKPKYSFRLLDPKY 280
QY 256 ----KSYNFRATATHWVEQPGVEARTLLKLYGIRFDILVTGQAGKFGLIPTAVTLGTGA 311
DB 281 DPASSGVNFRFAKYKINGTITRTTILIKAYGIRIDIVHGOAGKFSLIPTIINLATALTS 340
QY 312 LGVTFPFCOLLILLYVDREAHFYWRKYBEAKPKATANSVWRELALASQARLAECLERSS 371
DB 341 IGVSELCDEWILLTFMKNKLYSHKFKDKVTPKHPSSRWPVTALV-----L 388
QY 372 APAPTATAAGSOTQTPGWPCCPSSDT 396
DB 389 GQIPPPPSHYSDQPPSPSPSGEGPT 413

Search completed: November 21, 2004, 15:05:00
Job time : 41 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:41:41 ; Search time 156 Seconds  
(without alignments)  
931.316 Million cell updates/sec

Title: US-09-820-095b-2

Perfect score: 2226

Sequence: 1 MGSPARTGWLLDYKTEKW.....TFGWPCPSDTHLPHTHGSL 405

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358723299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2226	100.0	405	AAO15997	Human P2X
2	2203	99.0	431	AAW76434	Human p53
3	2203	99.0	431	ADD44744	Human Pro
4	2203	99.0	441	AAE01141	Human P2X
5	2199	98.8	431	ADP49183	Human P2X
6	1554	69.8	379	AAE01142	Rat Purin
7	1554	69.8	379	ADD44742	Rat Prote
8	1552	69.7	379	AAW76440	Human p53
9	1444.5	64.9	310	ADP29439	Human sec
10	892	40.1	455	AAW76439	Human p53
11	825.5	37.1	422	ADP07989	Novel p53
12	822	35.9	422	ADP07987	Novel p53
13	822	35.9	422	ADP49181	Human P2X
14	818	36.7	404	AAW47067	Human bra
15	816	36.7	388	AAW55035	Human bra
16	816	36.7	388	AAW55035	HPVRR ami
17	816	36.7	388	ADA09975	Human lig
18	816	35.7	388	ADH36222	Human pur
19	816	35.7	388	ADP49179	Human P2X
20	813	35.5	388	ADH4382	Amino aci
21	810	36.4	388	ADH1512	Human P2X
22	810	36.4	388	ADJ58917	Human 102
23	794	33.7	388	AAW04216	Rat super
24	791	35.5	388	AAW76438	Human p53
25	779	35.0	422	ABB97440	Novel hum

26	744	33.4	399	2	AAW04218	Human uri
27	744	33.4	399	7	ADA01480	Human pur
28	744	33.4	399	7	ADK52596	Hematolog
29	744	33.4	399	8	ADP49173	Human P2X
30	739	33.2	399	2	AAW04215	Rat vas d
31	739	33.2	399	2	AAW76435	Human p53
32	726	32.6	440	8	ADH22560	Human tra
33	720	32.3	472	2	AAW76436	Human p53
34	720	32.3	472	4	AAW28252	P2X 2. 2/
35	720	32.3	472	7	ADA09824	Rat P2X2
36	720	32.3	472	8	ADP49175	Rat P2X2
37	709	31.9	459	7	AAE38586	Human ion
38	709	31.9	471	3	AAW70274	Human pur
39	709	31.9	471	4	AAE84378	Amino aci
40	706.5	31.7	373	8	ADP29741	Human sec
41	704.5	31.6	397	6	ABR59707	Human ion
42	704.5	31.6	397	7	ADJ70396	Human hea
43	694	31.2	402	4	AAU87387	Novel cen
44	694	31.2	402	8	ADI54702	Novel hum
45	694	31.2	404	3	AAW70275	Human pur

#### ALIGNMENTS

##### RESULT 1

AAO15997  
ID AAO15997 standard; protein; 405 AA.  
XX AC AAO15997;

XX AC AAO15997;  
DT 20-FEB-2003 (first entry)  
XX DE Human P2X-like purinergic receptor G-protein coupled receptor.

XX DE Human; gene therapy; G-protein coupled receptor; chromosome 22;  
XX KW P2X-like purinergic receptor; anaplastic oligodendrogloma; GPCR;  
XX KW chronic lymphocytic leukaemia; lung carcinoma; colon carcinoma;  
XX KW brain carcinoma; drug screening; tissue typing; pharmacogenomic analysis;  
XX KW single nucleotide polymorphism; SNP.

XX OS Homo sapiens.

XX PN WO200279229-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US009545.

XX PR 29-MAR-2001; 2001US-00820095.

XX PA (PEKE ) PE CORP NY.

XX FI Wei M, Gong F, Di Francesco V, Beasley EM;

XX DR WPI; 2003-040648/03.

XX DR N-PSDB; AAL51048, AAL51049.

XX PT New peptides related to P2X-like purinergic receptor subfamily, useful for  
XX PT treating disorders associated with abnormal expression of protease in  
XX PT anaplastic oligodendrogloma, leukemia, carcinoma lung, or large cell  
XX PT lung carcinoma.

XX PS Claim 1; Fig 2A; 87pp; English.

XX CC The invention comprises the amino acid and coding sequence (located on  
XX CC chromosome 22) of a human G-protein coupled receptor (GPCR) which is  
XX CC related to the P2X-like purinergic receptor subfamily. The DNA and  
XX CC protein sequences of the invention are useful for treating: anaplastic  
XX CC oligodendrogloma; chronic lymphocytic leukaemia; lung carcinoma; colon  
XX CC carcinoma; and brain carcinoma. The DNA and protein sequences of the  
XX CC invention are also useful for drug screening assays, tissue typing and  
XX CC pharmacogenomic analysis. The present amino acid sequence represents the

CC human GPCR that is related to the P2X-like purinergic receptor subfamily  
 XX  
 SQ Sequence 405 AA;  
 Query Match 100.0%; Score 2226; DB 6; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-210;  
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGSPGATTGWLGDYKTEKALLAKGVQERDLEPQSIITKLKGVSVTQIKELGNRLWD 60  
 DB 1 MGSPGATTGWLGDYKTEKALLAKGVQERDLEPQSIITKLKGVSVTQIKELGNRLWD 60  
 QY 61 VADFVKPPQGENVFLVNTFLVTPAQVQRCPEHPSVPLANCWVDEDCPEGGTSHGV 120  
 DB 61 VADFVKPPQGENVFLVNTFLVTPAQVQRCPEHPSVPLANCWVDEDCPEGGTSHGV 120  
 QY 121 KTGQCVVNGTHRTCEIWSWCPVSGVVPSPRLAQAQNTLFIKNTVTSKFNFSKNA 180  
 DB 121 KTGQCVVNGTHRTCEIWSWCPVSGVVPSPRLAQAQNTLFIKNTVTSKFNFSKNA 180  
 QY 181 LETWDPYFKHCRVPEQSPYCPVFRIGDLVAKAGGFEDLALGGSGVIRVHWDCLDT 240  
 DB 181 LETWDPYFKHCRVPEQSPYCPVFRIGDLVAKAGGFEDLALGGSGVIRVHWDCLDT 240  
 QY 241 GDSGCWPHYSFQLOEKSYNFTATHHWEQPGVEARTLLKLYGIRFDILVTGQAGKGLIP 300  
 DB 241 GDSGCWPHYSFQLOEKSYNFTATHHWEQPGVEARTLLKLYGIRFDILVTGQAGKGLIP 300  
 QY 301 TAVTLGTGAALGVVTFPCDLLLVVDREAHFYWRTKYEEAKAPKATANSVWRELALASQ 360  
 DB 301 TAVTLGTGAALGVVTFPCDLLLVVDREAHFYWRTKYEEAKAPKATANSVWRELALASQ 360  
 QY 361 ARLAECLRRSSAPATAAGSQTPTGWPCCSSDTHLPTHSGSL 405  
 DB 361 ARLAECLRRSSAPATAAGSQTPTGWPCCSSDTHLPTHSGSL 405  
 RESULT 2  
 AAW76434  
 ID AAW76434 standard; protein; 431 AA.  
 XX  
 AC AAW76434;  
 XX  
 DT 28-APR-1999 (first entry)  
 XX  
 DE Human p53 regulated protein, P2XM.  
 XX  
 KW Human; p53 regulated protein; P2XM; ATP receptor; RP-2 protein;  
 KW thymocyte; apoptosis; tumour-suppressor gene; rhabdoid sarcoma;  
 KW anticancer drug.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO9842835-A1.  
 XX  
 PD 01-OCT-1998.  
 XX  
 PF 18-MAR-1998; 98WO-JP001146.  
 XX  
 PR 26-MAR-1997; 97JP-00093044.  
 XX  
 PA (SAKA) OTSUKA PHARM CO LTD.  
 XX  
 PI Tokino T, Nakamura Y;  
 XX  
 DR WPI; 1998-532006/45.  
 DR N-P5DB; AAV61832, AAV61833.  
 XX  
 PT Human gene P2XM whose transcription is induced by p53 - useful, e.g. for  
 PT diagnostic purposes and in development of new anticancer drugs.  
 XX  
 PS Claim 1; Page 23-27; 43pp; Japanese.

CC This sequence represents the P2XM protein of the invention. The protein  
 CC is significantly homologous to: (i) the P2X family of ATP receptors, and  
 CC (ii) Rp-2 protein which is expressed in thymocytes during apoptosis.  
 CC Transcription of the genes is specifically regulated by the tumour-  
 CC suppressor gene p53. The P2XM gene is specifically expressed in skeletal  
 CC muscle and has been localised to chromosome 22q11, an area where mutation  
 CC and sequence losses frequently occur in rhabdoid sarcomas. The genes may  
 CC be used for diagnostic purposes (e.g. by detecting changes occurring in  
 CC the gene in sarcomas), using probes and primers containing or derived  
 CC from all or part of the genes. The genes may further be used in the  
 CC development of new anticancer drugs  
 XX  
 SQ Sequence 431 AA;  
 Query Match 99.0%; Score 2203; DB 2; Length 431;  
 Best Local Similarity 94.0%; Pred. No. 7.8e-208;  
 Matches 405; Conservative 0; Mismatches 0; Indels 26; Gaps 1;  
 QY 1 MGSPGATTGWLGDYKTEK-----WALLAKKGYOERDLE 34  
 DB 1 MGSPGATTGWLGDYKTEKYVYVTRNRVGLQRLQLQIGIVVVGWALLAKKGYOERDLE 60  
 QY 35 POFISIITKLKGVSVTQIKELGNRLWDVADFVKPPQGENVFLVNTFLVTPAQVQRCPEH 94  
 DB 61 POFISIITKLKGVSVTQIKELGNRLWDVADFVKPPQGENVFLVNTFLVTPAQVQRCPEH 120  
 QY 95 PSVPLANCWVDEDCPEGGTSHGVKTGCQVFNHGTHTCEIWSWCPVSGVVPSPRL 154  
 DB 121 PSVPLANCWVDEDCPEGGTSHGVKTGCQVFNHGTHTCEIWSWCPVSGVVPSPRL 180  
 QY 155 AQAQNTLFIKNTVTSKFNFSKNALETWDPYFKHCRVPEQSPYCPVFRIGDLVAKA 214  
 DB 181 AQAQNTLFIKNTVTSKFNFSKNALETWDPYFKHCRVPEQSPYCPVFRIGDLVAKA 240  
 QY 215 GTFEDLALGGSGVIRVHWDCLDTGSGCWPHYSFQLOEKSYNFTATHHWEQPGVEA 274  
 DB 241 GTFEDLALGGSGVIRVHWDCLDTGSGCWPHYSFQLOEKSYNFTATHHWEQPGVEA 300  
 QY 275 RTLLKLYGIRFDILVTGQAGKGLIPTAVTLGTGAALGVVTFPCDLLLVVDREAHFYW 334  
 DB 301 RTLLKLYGIRFDILVTGQAGKGLIPTAVTLGTGAALGVVTFPCDLLLVVDREAHFYW 360  
 QY 335 RTKYEEAKAPKATANSVWRELALASQARLAECLRRSSAPATAAGSQTPTGWPCCSS 394  
 DB 361 RTKYEEAKAPKATANSVWRELALASQARLAECLRRSSAPATAAGSQTPTGWPCCSS 420  
 QY 395 DTHLPTHSGSL 405  
 DB 421 DTHLPTHSGSL 431  
 RESULT 3  
 ADD44744  
 ID ADD44744 standard; protein; 431 AA.  
 XX  
 AC ADD44744;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human Protein NP\_005437, SEQ ID NO 10173.  
 XX  
 KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.



PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX (GEO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
XX  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
PI  
DR WPI; 2003-268312/26.  
DR GENBANK; NP\_005437.  
XX  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX  
PS Claim 1; Page; 1017pp; English.  
XX  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 431 AA;  
  
Query Match 99.0%; Score 2203; DB 7; Length 431;  
Best Local Similarity 94.0%; Pred. No. 7.8e-208;  
Matches 405; Conservative 0; Mismatches 0; Indels 26; Gaps 1;  
  
QY 1 MGSPGATTGGLLDYKTER-----WALLAKGYQERDLE 34  
DB 1 MGSPGATTGGLLDYKTERVWTRNVRGALQRLQFGIVVVGWALLAKGYQERDLE 60  
  
QY 35 POFSIITKLKGVSVTQIKELGNRLWDVADVPQGENVFFLVTLVTPAQVQRCPEH 94  
DB 61 POFSIITKLKGVSVTQIKELGNRLWDVADVPQGENVFFLVTLVTPAQVQRCPEH 120  
  
QY 95 PSVPLANCWVDDCEGEGCTSHGKVTQCQVFNVTGRTCEIMGWCVPESGWPSRPLL 154  
DB 121 PSVPLANCWVDDCEGEGCTSHGKVTQCQVFNVTGRTCEIMGWCVPESGWPSRPLL 180  
  
QY 155 AQAQNTFLIKNTVTFKFNFSKNALETWDTYFKHCRYEPQFPYCPFRIGDLVAKA 214  
DB 181 AQAQNTFLIKNTVTFKFNFSKNALETWDTYFKHCRYEPQFPYCPFRIGDLVAKA 240  
  
QY 215 GGTTFEDALLGGSVGIRVHWDCLDTGDSGCWPHYSFQLEKSYNFRATTHWEOPGVEA 274  
DB 241 GGTTFEDALLGGSVGIRVHWDCLDTGDSGCWPHYSFQLEKSYNFRATTHWEOPGVEA 300  
  
QY 275 RTLLKLYGIRFDILVTGQAGKGLIPTAVTLTGTAANLGVVTFPCDLLLLLYVDREAHFYW 334

DB 301 RTLLKLYGIRFDILVTGQAGKGLIPTAVTLTGTAANLGVVTFPCDLLLLLYVDREAHFYW 360  
QY 335 RTKYEEAKAPKATANSVWRELALASQARLAECRLRRSSAPAPTATAAGSQTTPGWPCCSS 394  
DB 361 RTKYEEAKAPKATANSVWRELALASQARLAECRLRRSSAPAPTATAAGSQTTPGWPCCSS 420  
QY 395 DTHLPTHSGSL 405  
DB 421 DTHLPTHSGSL 431  
  
RESULT 4  
AAE01141  
ID AAE01141 standard; protein; 441 AA.  
XX  
XX AAE01141;  
AC  
DT 17-JUL-2001 (first entry)  
DE Human purinergic receptor P2X6 protein.  
XX  
XX Human; purinergic receptor; P2X6; antidepressant; vulnerary; hypotensive;  
KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;  
KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;  
KW neuromuscular disease; reproductive system disorder; hypertension;  
KW peripheral vascular disease; immune system disorder; chronic bronchitis;  
KW irritable bowel disorder; premature ejaculation; asthma.  
XX  
OS Homo sapiens.  
XX  
PN USG214581-B1.  
XX  
PD 10-APR-2001.  
XX  
PF 13-NOV-1998; 98US-00191136.  
XX  
PR 16-JAN-1998; 98US-00008185.  
PR 16-JAN-1998; 98US-00008526.  
PR 16-JAN-1998; 98US-0071298P.  
XX 16-JAN-1998; 98US-0071669P.  
PA (ABBO) ABBOTT LAB.  
XX  
PI Lynch KJ, BURGARD EC, Van Biesen T;  
XX  
DR WPI; 2001-315459/33.  
XX N-PSDB; RAD04978.  
PT Novel isolated polynucleotide encoding human purinergic P2X3 receptor  
PT polypeptide useful for identifying potentially therapeutic compounds that  
PT modulate or otherwise interact with P2X containing receptors.  
XX  
XX Example 9; Fig 9; 53pp; English.  
XX  
XX The present sequence is human P2X6 purinergic receptor protein. P2X  
CC receptors are ligand-gated ion channels while P2Y receptors operate  
CC generally through a G-protein coupled system. P2X purinoreceptor drugs  
CC are potential therapeutic agents in several disorders including central  
CC nervous system or peripheral nervous system conditions, e.g., epilepsy,  
CC pain, depression, neurodegenerative disorders, disorders of the skeletal  
CC muscle such as neuromuscular diseases, disorders of the reproductive  
CC system, asthma, peripheral vascular disease, hypertension, immune system  
CC disorders, irritable bowel disorder, premature ejaculation, cystic  
CC fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the activity  
CC of extracellular nucleotide triphosphates to regulate chloride secretion  
CC in human airway epithelia  
XX  
SQ Sequence 441 AA;  
  
Query Match 99.0%; Score 2203; DB 4; Length 441;  
Best Local Similarity 94.0%; Pred. No. 8.1e-208;  
Matches 405; Conservative 0; Mismatches 0; Indels 26; Gaps 1;

CC studying the signal transduction pathways related to olfactory signaling.  
 CC The compositions are also useful as reagents in micro-arrays or as  
 CC reagents to probe or analyze existing micro-arrays. This sequence  
 CC corresponds to the human P2X6 protein sequence.

XX  
 SQ Sequence 431 AA;

Query Match 98.8%; Score 2199; DB 8; Length 431;  
 Best Local Similarity 93.7%; Pred. No. 1.9e-207;  
 Matches 404; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

QY 1 MGSPGATTGWLGLDYKTEK-----WALLAKGYQERDLE 34  
 DB 11 MGSPGATTGWLGLDYKTEKYVTRNRVGAQLQLQFGIVVVVWALLAKGYQERDLE 70  
 QY 35 PQFSIITKLKGVSVTQIKELGNRLMDVADFVKPPQGENVFFLVTNFLTTPAQVQGRCPHEH 94  
 DB 71 PQFSIITKLKGVSVTQIKELGNRLMDVADFVKPPQGENVFFLVTNFLTTPAQVQGRCPHEH 130  
 QY 95 PSVPLANCWVDEDCPEGEGGTHSHGVKTGCQVFNHGTCTCEIWSWCPVSGVPSRPLL 154  
 DB 131 PSVPLANCWVDEDCPEGEGGTHSHGVKTGCQVFNHGTCTCEIWSWCPVSGVPSRPLL 190  
 QY 155 AQAQNFLLFKNTVTFKSNFNSKNALETWDPYFKHCRYPQSPYCPVFRIGDLVAKA 214  
 DB 191 AQAQNFLLFKNTVTFKSNFNSKNALETWDPYFKHCRYPQSPYCPVFRIGDLVAKA 250  
 QY 215 GGTFFEDLALLGGSGVIRVHWDCCDLTGDSGCWPHYSFQLOEKSYNFRATATHWQPGVEA 274  
 DB 251 GGTFFEDLALLGGSGVIRVHWDCCDLTGDSGCWPHYSFQLOEKSYNFRATATHWQPGVEA 310  
 QY 275 RTLLKLYGIRFDILVTGQAGKFGELIPTAVTLGTGAWLGVVTFDCDLLLLYVDREAHFYW 334  
 DB 311 RTLLKLYGIRFDILVTGQAGKFGELIPTAVTLGTGAWLGVVTFDCDLLLLYVDREAHFYW 370  
 QY 335 RTKYEEAKAPKATANSVWRELALASQARLAECRLRSSAPAPTATAAGSQTPTPGWPCPSS 394  
 DB 371 RTKYEEAKAPKATANSVWRELALASQARLAECRLRSSAPAPTATAAGSQTPTPGWPCPSS 430  
 QY 395 DTHLPTHSGSL 405  
 DB 431 DTHLPTHSGSL 441

## RESULT 5

ADP49183  
 ID ADP49183 standard; protein; 431 AA.  
 XX  
 AC ADP49183;  
 XX  
 DT 26-AUG-2004 (first entry)  
 XX  
 DE Human P2X6 purinergic receptor protein sequence for odour modulation.  
 XX  
 DE odour sensitivity; P2X purinergic receptor; P2Y purinergic receptor;  
 KW signal transduction pathway; olfactory signalling; micro-array.  
 KW  
 XX Homo sapiens.  
 XX  
 PN WO2004047749-A2.  
 XX  
 PD 10-JUN-2004.  
 XX  
 PF 21-NOV-2003; 2003WO-US037389.  
 XX  
 PR 21-NOV-2002; 2002US-0428140P.  
 XX  
 PA (UTAH ) UNIV UTAH RES FOUND.  
 XX  
 XX Lucero M, Hegg C;  
 XX  
 XX WPI; 2004-460642/43.  
 XX  
 XX Modulating odor sensitivity in a subject, comprises administering a  
 PT composition comprising an agonist or antagonist of P2X or P2Y purinergic  
 PT receptor to the subject.  
 XX  
 XX Disclosure; SEQ ID NO 12; 108pp; English.  
 XX  
 XX The invention relates to a method of modulating (M1) odour sensitivity in  
 CC a subject, by administering a composition which is an agonist or  
 CC antagonist of a P2X or P2Y purinergic receptor. (M1) is useful for  
 CC modulating odour sensitivity in a subject (claimed). The compositions  
 CC used for modulating odour sensitivity in a subject are useful for

CC studying the signal transduction pathways related to olfactory signaling.  
 CC The compositions are also useful as reagents in micro-arrays or as  
 CC reagents to probe or analyze existing micro-arrays. This sequence  
 CC corresponds to the human P2X6 protein sequence.

XX  
 SQ Sequence 431 AA;

Query Match 98.8%; Score 2199; DB 8; Length 431;  
 Best Local Similarity 93.7%; Pred. No. 1.9e-207;  
 Matches 404; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

QY 1 MGSPGATTGWLGLDYKTEK-----WALLAKGYQERDLE 34  
 DB 1 MGSPGATTGWLGLDYKTEKYVTRNRVGAQLQLQFGIVVVVWALLAKGYQERDLE 60  
 QY 35 PQFSIITKLKGVSVTQIKELGNRLMDVADFVKPPQGENVFFLVTNFLTTPAQVQGRCPHEH 94  
 DB 61 PQFSIITKLKGVSVTQIKELGNRLMDVADFVKPPQGENVFFLVTNFLTTPAQVQGRCPHEH 120  
 QY 95 PSVPLANCWVDEDCPEGEGGTHSHGVKTGCQVFNHGTCTCEIWSWCPVSGVPSRPLL 154  
 DB 121 PSVPLANCWVDEDCPEGEGGTHSHGVKTGCQVFNHGTCTCEIWSWCPVSGVPSRPLL 180  
 QY 155 AQAQNFLLFKNTVTFKSNFNSKNALETWDPYFKHCRYPQSPYCPVFRIGDLVAKA 214  
 DB 181 AQAQNFLLFKNTVTFKSNFNSKNALETWDPYFKHCRYPQSPYCPVFRIGDLVAKA 240  
 QY 215 GGTFFEDLALLGGSGVIRVHWDCCDLTGDSGCWPHYSFQLOEKSYNFRATATHWQPGVEA 274  
 DB 241 GGTFFEDLALLGGSGVIRVHWDCCDLTGDSGCWPHYSFQLOEKSYNFRATATHWQPGVEA 300  
 QY 275 RTLLKLYGIRFDILVTGQAGKFGELIPTAVTLGTGAWLGVVTFDCDLLLLYVDREAHFYW 334  
 DB 301 RTLLKLYGIRFDILVTGQAGKFGELIPTAVTLGTGAWLGVVTFDCDLLLLYVDREAHFYW 360  
 QY 335 RTKYEEAKAPKATANSVWRELALASQARLAECRLRSSAPAPTATAAGSQTPTPGWPCPSS 394  
 DB 361 RTKYEEAKAPKATANSVWRELALASQARLAECRLRSSAPAPTATAAGSQTPTPGWPCPSS 420  
 QY 395 DTHLPTHSGSL 405  
 DB 421 DTHLPTHSGSL 431

## RESULT 6

AAE01142  
 ID AAE01142 standard; protein; 379 AA.  
 XX  
 AC AAE01142;  
 XX  
 DT 17-JUL-2001 (first entry)  
 XX  
 DE Rat purinergic receptor P2X6 protein.  
 XX  
 XX Rat; purinergic receptor; P2X6; antidepressant; vulnerary; hypotensive;  
 KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;  
 KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;  
 KW neuromuscular disease; reproductive system disorder; hypertension;  
 KW peripheral vascular disease; immune system disorder; chronic bronchitis;  
 KW irritable bowel disorder; premature ejaculation; asthma.  
 XX  
 XX Rattus sp.  
 OS  
 XX US6214581-B1.  
 PN  
 XX 10-APR-2001.  
 PD  
 XX 13-NOV-1998; 98US-00191136.  
 PF  
 XX 16-JAN-1998; 98US-00008185.  
 PR 16-JAN-1998; 98US-00008526.  
 PR 16-JAN-1998; 98US-0071298P.  
 PR 16-JAN-1998; 98US-0071669P.

XX PA (ABBO ) ABBOTT LAB.  
 XX PI Lynch KJ, Burgard EC, Van Biesen T;  
 XX DR WPI; 2001-315459/33.  
 XX PT Novel isolated polynucleotide encoding human purinergic P2X3 receptor  
 XX PT polypeptide useful for identifying potentially therapeutic compounds that  
 XX PT modulate or otherwise interact with P2X containing receptors.  
 XX PS Example 9; Fig 9; 53pp; English.  
 XX CC The present sequence is rat P2X6 purinergic receptor protein. P2X  
 CC receptors are ligand-gated ion channels while P2X receptors operate  
 CC generally through a G-protein coupled system. P2X purinoreceptor drugs  
 CC are potential therapeutic agents in several disorders including central  
 CC nervous system or peripheral nervous system conditions, e.g., epilepsy,  
 CC pain, depression, neurodegenerative disorders, disorders of the skeletal  
 CC muscle such as neuromuscular diseases, disorders of the reproductive  
 CC system, asthma, peripheral vascular disease, hypertension, immune system  
 CC disorders, irritable bowel disorder, premature ejaculation, cystic  
 CC fibrosis and chronic bronchitis. P2X purinoreceptors mediate the activity  
 CC of extracellular nucleotide triphosphates to regulate chloride secretion  
 CC in human airway epithelia  
 XX SQ Sequence 379 AA;

Query Match 69.8%; Score 1554; DB 4; Length 379;  
 Best Local Similarity 75.7%; Pred. No. 5.8e-144;  
 Matches 281; Conservative 26; Mismatches 38; Indels 26; Gaps 1;  
 QY 6 ATTGGLLDYKTEK-----WALLAKKGQERDLBPQFSI 39  
 DB 8 ALVSWGFLDYKTEKYVMTRNCWVGISQRLQLQGVVYVIGWALLAKKGQERDMDPQISV 67  
 QY 40 ITKLKGVSVTQIKELGNRLNDVADFVKPPQGENVFLVNTVTPAQVQGRCPHEPSVPL 99  
 DB 68 ITKLKGVSVTQVKELEKRLNDVADFVRPSQGENVFLVNTVTPAQVQGRCPHEPSVPL 127  
 QY 100 ANCWDEDCPEGEGGTHSHGVKTCQCVFNGTHTCTEISWCPVSVSRPPLAQON 159  
 DB 128 ANCWDEDCPEGEGGYSHGKTCQCVAFNGTHTCTEISWCPVSVSAVPRKPLLAQKN 187  
 QY 160 FTLFIKNTVTFKFNFSKNALETWDPVTFKHCRVEPQSPYCPVFRIGDLVAKAGTPE 219  
 DB 188 FTLFIKNTVTFKFNFSRNTALDWDNTYFKYCLVDSUSSPCVFRIGDLVAMVGGDFE 247  
 QY 220 DLALLGSGVIRVHWDCLDLDGSGCWPHYSFQLQEKSYNFRATATHWEQPGVEARTLLK 279  
 DB 248 DLALLGAGVINIHWDCNLDTKGDCSPQYSFQLQERGINFRATANYWAAAGVESRSLJK 307  
 QY 280 LYGIRFDILVTGQAGKFLIPTAVTLGTGAAMLGWTFPCDLLLLYVDREAHFYWRTRYE 339  
 DB 308 LYGIRFDILVTGQAGKFLIPTAITVTGTGAAMLGWTFCLDLLLLYVDREAGFYWRTRYE 367  
 QY 340 EAKAPKATANS 350  
 DB 368 EAKAPKATNS 378

RESULT 7  
 ADD44742  
 ID ADD44742 standard; protein; 379 AA.  
 XX AC ADD44742;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Rat Protein P51579, SEQ ID NO 10171.  
 XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.  
 XX PN WO2003016475-A2.  
 XX PD 27-FEB-2003.  
 XX PF 14-AUG-2002; 2002WO-US025765.  
 XX PR 14-AUG-2001; 2001US-0312147P.  
 XX PR 01-NOV-2001; 2001US-0346382P.  
 XX PR 26-NOV-2001; 2001US-0333347P.  
 XX PA (GEO ) GEN HOSPITAL CORP.  
 XX PA (FARB ) BAYER AG.  
 XX PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 XX DR GENBANK; P51579.  
 XX PT New composition comprising two or more isolated polypeptides, useful for  
 XX PT preparing a medicament for treating pain in an animal.  
 XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 379 AA;

Query Match 69.8%; Score 1554; DB 7; Length 379;  
 Best Local Similarity 75.7%; Pred. No. 5.8e-144;  
 Matches 281; Conservative 26; Mismatches 38; Indels 26; Gaps 1;

QY 6 ATTGGLLDYKTEK-----WALLAKKGQERDLBPQFSI 39  
 DB 8 ALVSWGFLDYKTEKYVMTRNCWVGISQRLQLQGVVYVIGWALLAKKGQERDMDPQISV 67  
 QY 40 ITKLKGVSVTQIKELGNRLNDVADFVKPPQGENVFLVNTVTPAQVQGRCPHEPSVPL 99  
 DB 68 ITKLKGVSVTQVKELEKRLNDVADFVRPSQGENVFLVNTVTPAQVQGRCPHEPSVPL 127  
 QY 100 ANCWDEDCPEGEGGTHSHGVKTCQCVFNGTHTCTEISWCPVSVSRPPLAQON 159  
 DB 128 ANCWDEDCPEGEGGYSHGKTCQCVAFNGTHTCTEISWCPVSVSAVPRKPLLAQKN 187  
 QY 160 FTLFIKNTVTFKFNFSKNALETWDPVTFKHCRVEPQSPYCPVFRIGDLVAKAGTPE 219

Db 188 FTLFIKNTVTNFKNFRTNALDWTNDTFFYKCYDLSLSSPYCFVFRIGDLVAMTGGDFE 247  
 Qy 220 DLALLGGVGIRVHWDCLDGTGDCSWPHYSFQIQEKSYNFRTATHWQFQVEARTLK 279  
 Db 248 DLALLGGVGINIHWCNLDTKGSDCSPOYSFQIQEYGNFRTANYWMAAGVBSRLK 307  
 Qy 280 LYGIRFDILVTGQAGKFLIPTAVTLGTGAALWGVVTFPCDLLLLLYVDREAHFYWRKYE 339  
 Db 308 LYGIRFDILVTGQAGKFLIPTAITVGTGAALWGVVTFPCDLLLLLYVDREAGFYWRKYE 367  
 Qy 340 EAKAPKATANS 350  
 Db 368 EAKAPKATANS 378

RESULT 8  
 AAW76440  
 ID AAW76440 standard; protein; 379 AA.  
 XX  
 AC AAW76440;  
 DT 28-APR-1999 (first entry)  
 XX  
 DE Human p53 regulated protein, P2X6.  
 XX  
 KW Human; p53 regulated protein; P2XM; ATP receptor; RP-2 protein;  
 KW thymocyte; apoptosis; tumour-suppressor gene; thadoid sarcoma;  
 KW anticancer drug.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9842835-A1.  
 XX  
 PD 01-OCT-1998.  
 XX  
 PF 18-MAR-1998; 98WO-JP001146.  
 XX  
 PR 26-MAR-1997; 97JP-00093044.  
 XX  
 PA (SAXA ) OTSUKA PHARM CO LTD.  
 XX  
 PI Tokino T, Nakamura Y;  
 XX  
 DR WPI: 1998-532006/45.  
 XX  
 PT Human gene P2XM whose transcription is induced by p53 - useful, e.g. for  
 PT diagnostic purposes and in development of new anticancer drugs.  
 XX  
 PS Disclosure; Fig 5-6; 43pp; Japanese.  
 XX  
 CC This sequence represents the human P2X6 protein. The invention relates to  
 CC the P2XM protein, which is significantly homologous to: (i) the P2X  
 CC family of ATP receptors, and (ii) RP-2 protein which is expressed in  
 CC thymocytes during apoptosis. Transcription of the genes is specifically  
 CC regulated by the tumour-suppressor gene p53. The P2XM gene is  
 CC specifically expressed in skeletal muscle and has been localised to  
 CC chromosome 22q11, an area where mutation and sequence losses frequently  
 CC occur in thadoid sarcomas. The genes may be used for diagnostic purposes  
 CC (e.g. by detecting changes occurring in the gene in sarcomas), using  
 CC probes and primers containing or derived from all or part of the genes.  
 CC The genes may further be used in the development of new anticancer drugs  
 XX  
 SQ Sequence 379 AA;

Query Match 69.7%; Score 1552; DB 2; Length 379;  
 Best Local Similarity 75.7%; Pred. No. 9.1e-144;  
 Matches 281; Conservative 26; Mismatches 38; Indels 26; Gaps 1;  
 Qy 6 ATTGWLGLDYKTEK-----WALLAKGYQERDLEPQPSI 39  
 Db 8 ALVSWGFLDYKTEKYMTNVCWVGISORLLQLGVVVVIGWALLAKGYQEWMDPQISV 67

Qy 40 ITKLKGVSVTQIKELGNRLWDVADFVKPQGENVFFLVNTFLVTPAQVQGRCPHEHPSVPL 99  
 Db 68 ITKLKGVSVTQIKELKRLWDVADFVRPSQGENVFFLVNTFLVTPAQVQGRCPHEHPSVPL 127  
 Qy 100 ANCWVDEDCPEGEGETHSHGVKTGCVVFNPGHRTCEIWSWCPVSVVPSRPLLAQON 159  
 Db 128 ANCWVDEDCPEGEGETHSHGVKTGCVVFNPGHRTCEIWSWCPVSVVPSRPLLAQON 187  
 Qy 160 FTLFIKNTVTNFKNFRTNALDWTNDTFFYKCYDLSLSSPYCFVFRIGDLVAMTGGDFE 219  
 Db 188 FTLFIKNTVTNFKNFRTNALDWTNDTFFYKCYDLSLSSPYCFVFRIGDLVAMTGGDFE 247  
 Qy 220 DLALLGGVGIRVHWDCLDGTGDCSWPHYSFQIQEKSYNFRTATHWQFQVEARTLK 279  
 Db 248 DLALLGGVGINIHWCNLDTKGSDCSPOYSFQIQEYGNFRTANYWMAAGVBSRLK 307  
 Qy 280 LYGIRFDILVTGQAGKFLIPTAVTLGTGAALWGVVTFPCDLLLLLYVDREAHFYWRKYE 339  
 Db 308 LYGIRFDILVTGQAGKFLIPTAITVGTGAALWGVVTFPCDLLLLLYVDREAGFYWRKYE 367  
 Qy 340 EAKAPKATANS 350  
 Db 368 EAKAPKATANS 378  
 RESULT 9  
 ADP29439  
 ID ADP29439 standard; protein; 310 AA.  
 XX  
 AC ADP29439;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #206.  
 XX  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KW cancer; inflammatory; immune; human secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406653P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410946P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.

D	b	1	WGLLDYKTEKYVMTRNRVRGALQRLLOFGIVVYVVGWALLAKKGYGQRDLPEQFSIITKL	60
Q	y	44	KGVSVTQIKELGNRLMDVADFPKPPQGENVFFLVTNFLVTPAQVQGRCPHPSPVPLANCW	103
D	b	61	KGVSVTQIKELGNRLMDVADFPKPPQGENVFFLVTNFLVTPAQVQGRCPHPSPVPLANCW	120
Q	y	104	VDECPGEGGTHSHGVKTCQCVVFNQTHRTCEIWSWCPVESGVVPSRPLLAQAQNFTLF	163
D	b	121	VDECPGEGGTHSHGVKTCQCVVFNQTHRTCEIWSWCPVESGVVPS-PLLAQAQNFTLF	179
Q	y	164	IKNTVTSKFNFSKNALETWDTPTYPKHCRYEYPQSPYCPVFRIGDLIVAKAGGTFEDLAL	223
D	b	180	IKNTVTSKFP-----VFRIGDLIVAKAGGTFEDLAL	209
Q	y	224	LGSSVGIRVHWDCLDTGDSGCWPHYSFQLQEKSYNFRATHTHWEQPGVEARTLLKLYGI	283
D	b	210	LGSSVGIRVHWDCLDTGDSGCWPHYSFQLQEKSYNFRATHTHWEQPGVEARTLLKLYGI	269
Q	y	284	RFDLILVTGQAGKFLIPTAVTLGTGAWLGV	314
D	b	270	RFDLILVTGQAGKFLIPTAVTLGTGAWLGV	300
RESULT 10				
AAW76439				
ID	AAW76439 standard; protein; 455 AA.			
XX	AC	AAW76439;		
XX	AC			
DT	28-APR-1999 (first entry)			
XX	XX			
DE	Human p53 regulated protein, P2X5.			
XX	XX			
KW	Human; p53 regulated protein; P2XM; ATP receptor; RP-2 protein;			
KW	thymocyte; apoptosis; tumour-suppressor gene; rhabdoid sarcoma;			
KW	anticancer drug.			
XX	XX			
OS	Homo sapiens.			
XX	XX			
PN	PN	W09842835-A1.		
PD	PD	01-OCT-1998.		
XX	XX	18-MAR-1998; 98WO-JP001146.		
XX	XX	26-MAR-1997; 97JP-00093044.		
PR	XX	(SAKA ) OTSUKA PHARM CO LTD.		
PA	XX			
XX	XX	Tokino T, Nakamura Y;		
PI	PI	WPI; 1998-532006/45.		
DR	DR			
XX	XX	Human gene P2XM whose transcription is induced by p53 - useful, e.g. for		
PT	PT	diagnostic purposes and in development of new anticancer drugs.		
XX	XX	Disclosure; Fig 5-6; 43pp; Japanese.		
XX	XX			
CC	CC	This sequence represents the human P2X5 protein. The invention relates to		
CC	CC	the P2XM protein, which is significantly homologous to: (i) the P2X		
CC	CC	family of ATP receptors, and (ii) RP-2 protein which is expressed in		
CC	CC	thymocytes during apoptosis. Transcription of the genes is specifically		
CC	CC	regulated by the tumour-suppressor gene p53. The P2XM gene is		
CC	CC	specifically expressed in skeletal muscle and has been localised to		
CC	CC	chromosome 22q11, an area where mutation and sequence losses frequently		
CC	CC	occur in rhabdoid sarcomas. The genes may be used for diagnostic purposes		
CC	CC	(e.g. by detecting changes occurring in the gene in sarcomas), using		
CC	CC	probes and primers containing or derived from all or part of the genes.		
CC	CC	The genes may further be used in the development of new anticancer drugs		
XX	SQ	Sequence 455 AA;		

Query Match 40.1%; Score 892; DB 2; Length 455;

Best Local Similarity 46.6%; Pred. No. 1.3e-78;  
Matches 174; Conservative 55; Mismatches 108; Indels 36; Gaps 5;

QY 12 LLDYKTEK-----WALLAKKGYOERDLEPOFSIITLKG 45  
DB 13 LFDYKTAFFVAKSKKGLLRVLQILILLVLLWFLIKSYODITSLQSAVTVKVG 72  
QY 46 VSVTQIKELGNRLMDVADFKVPQGENVFFLVTFNPLVTPAQVQRCPEHPSVPLANCWD 105  
DB 73 VAYTNTMLGERLMDVADFKVPSQGENVFFLVTFNPLVTPNQRCICAEREGIPDGECS 132  
QY 106 EDCPEGEGTHSHGVKTCQCV-VFNGHTRTCEIWSVPSVPSRPLLAQONETLFI 164  
DB 133 DCHAGESVWAGHGLTKRCLRVGNSTRGTCEIFAWCPVETKSMPTDPLKDAESFTISI 192  
QY 165 KNTVTFKSNFNSKNALETWPTFKHCRYPQSPYCPVFRIGDLVAKAGGTFFEDIAL 224  
DB 193 KNFIRPKFNFSKANVLETDNKHFLKTHFS-STNLYCPIPLGSIIVRWAGADFODIALK 251  
QY 225 GSGVIRVHWDCLDITGSGGWPHYSF-QLQEK-----SYNFRATHWQPCVEART 276  
DB 252 GGVIGIYIEWDCDLDAKAAKCNPHYFNRLDNKTHSISSGYNFRFARYVRDPNGVEFRD 311  
QY 277 LKLYGIRFDILVTQAGKGLIPTAVTLGTGAALWGVVTFCDLLLYVDREAHFYWRT 336  
DB 312 LMKAYGIRFDVIVNGKAKFSIPIPTVINIGSLALMGAGAFCDLVLIYLRKSEFYRDK 371  
QY 337 KYBEAKAPKATAN 349  
DB 372 KPEKVRGQKEDAN 384

RESULT 11  
ADE07989  
ID ADE07989 standard; protein; 422 AA.  
AC ADE07989;  
XX  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Novel protein (useful for identifying genetic disorders) #144.  
XX  
KW novel gene; novel protein; tissue marker; molecular weight marker;  
KW chromosome marker; genetic disorder.  
XX  
OS Unidentified.  
XX  
PN WO2003054152-A2.  
XX  
PD 03-JUL-2003.  
XX  
PF 10-DEC-2002; 2002WO-US039555.  
XX  
PR 10-DEC-2001; 2001US-0339739P.  
PR 11-DEC-2001; 2001US-0339453P.  
PR 14-MAR-2002; 2002US-0365091P.  
PR 14-MAR-2002; 2002US-0365384P.  
PR 12-APR-2002; 2002US-0372381P.  
PR 12-APR-2002; 2002US-0372615P.  
PR 22-APR-2002; 2002US-00128558.  
PR 24-APR-2002; 2002US-0376043P.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
XX  
DR WPI; 2003-569235/53.  
DR N-PSDB; ADE07078.  
XX  
XX New polynucleotides, useful for expressing recombinant proteins for  
PT analysis, characterization or therapeutic use, or as markers for tissues

PT in which the corresponding protein is preferentially expressed.

PS Claim 20; SEQ ID NO 1055; 1177pp; English.  
XX  
XX The invention comprises the amino acid and coding sequences of novel  
CC proteins. The DNA and protein sequences of the invention are useful as:  
CC markers for tissues in which the corresponding protein is preferentially  
CC expressed; as molecular weight markers on gels; as chromosome markers or  
CC tags; to identify chromosomes or to map related gene positions; and to  
CC compare with endogenous DNA sequences in patients to identify potential  
CC genetic disorders. The present amino acid sequence represents a protein  
CC of the invention.  
XX  
SQ Sequence 422 AA;

Query Match 37.1%; Score 825.5; DB 7; Length 422;  
Best Local Similarity 42.7%; Pred. No. 4.1e-72;  
Matches 166; Conservative 63; Mismatches 125; Indels 35; Gaps 7;

QY 20 WALLAKKGYOERDLEPOFSIITLKGVSVTQIKELGNRLMDVADFKVPQGENVFFLVTN 79  
DB 46 WFFVWEKGQETD-SVSSVTIKVGVAVTNTSKLGFINDVADIVIPAQENSILFVMTN 104  
QY 80 FLVTPAQVQRCPEHPSVPLANCWDEDCPEGEGTHSHGVKTCQCVFNGHTRTCEIWS 139  
DB 105 VILTMNQTLGCLPEIPDATTV-CKSDASCTAGTSAGTSHNGVSTGRCVAFNGSVKTCVAA 163  
QY 140 WCPVESGV-VPSRPLLAQONETLFIKNTVTFKSNFNSKNALETWPTFKHCRYPEQF 198  
DB 164 WCPVEDDTHVPOPAFLKAAENFTLLVKNINWPKNFNKRNLFPNITTYLKSCIYDAKT 223  
QY 199 SPYCPVFRIGDLVAKAGGTFFEDIALIGSGVIRVHWDCLDITGSGGWPHYSFQJQE--- 255  
DB 224 DPFCPIFRIGKIVENAGHSFQDMAVEGGIMGIVQVWDCNLDRAASLCLPRYSFRELDRD 283  
QY 256 -----KSNFRATHWQPCVEARTLLKLYGIRFDILVTQAGKGLIPTAVTLGTGA 309  
DB 284 VEHNVSFGYNFRFARYRDLAGNEQRTLIKAYGIRFDIIVFGKAGKFDIIPITMINIGSGL 343  
QY 310 AWLGVTFFCDLLLYVDREAHFYWRTKYEEAKAPKATANSVWRELALASQARLAECLE 369  
DB 344 ALLGMATVLCDIILVLYCMKRLYREKKYVYEDYEQVGPS-W----- 385  
QY 370 SSAPAPTATAA---GSQTPTPCWPCPSSD 395  
DB 386 --PPAGTGLSLGCGSPGRGPVWGGPSAE 412

RESULT 12  
ADF76897  
ID ADF76897 standard; protein; 422 AA.  
XX  
XX ADF76897;  
AC  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Novel human secreted and transmembrane protein SeqID 572.  
XX  
KW human; PRO; membrane bound protein; membrane bound receptor;  
KW cell proliferation; cell migration; cell differentiation;  
KW mitogenic factor; survival factor; cytotoxic factor;  
KW differentiation factor; neuropeptide; hormone; cell receptor;  
KW receptor-ligand interaction; cytostatic; chondrocyte; tumour.  
XX  
XX Homo sapiens.  
XX  
XX WO2003072035-A2.  
XX  
XX 04-SEP-2003.  
XX  
XX 21-FEB-2003; 2003WO-US005241.  
XX  
XX 22-FEB-2002; 2002US-0359461P.  
XX  
XX

XX PA (GETH ) GENENTECH INC.  
 XX PI Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;  
 XX PI Williams PM, Wood WI, Wu TD;  
 XX DR N-PSDB; ADF76896.  
 XX DR WPI; 2003-721702/68.  
 XX PT New PRO polypeptides, useful for diagnosing and treating an immune  
 PT related disorder, e.g. systemic lupus erythematosus, rheumatoid  
 PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or  
 PT diabetes mellitus.  
 XX PS Claim 10; SEQ ID NO 572; 918pp; English.  
 XX CC This invention relates to novel nucleic acids encoding human PRO secreted  
 CC and transmembrane proteins. Extracellular proteins play important roles  
 CC in the formation, differentiation and maintenance of multicellular  
 CC organisms. The fate of many individual cells (for example proliferation,  
 CC migration or differentiation) is typically governed by information  
 CC received from other cells and the immediate environment. The information  
 CC is often transmitted by secreted polypeptides (for example mitogenic  
 CC factors, survival factors, cytotoxic factors, differentiation factors,  
 CC neuropeptides and hormones) which are received and interpreted by diverse  
 CC cell receptors or membrane bound proteins. These membrane bound proteins  
 CC and receptors may be of use as pharmaceutical and diagnostic agents, such  
 CC as in the blocking of receptor-ligand interactions. The current invention  
 CC provides the amino acid sequences of novel human membrane bound receptors  
 CC and proteins, along with the cDNA sequences encoding them. The novel  
 CC proteins of the invention may have cytosolic activities through the  
 CC stimulation of chondrocytes. The nucleic acids of the invention may be  
 CC useful for the manufacture of a medicament for diagnosing or treating a  
 CC tumour in a mammal. In addition, they may be useful for measuring or  
 CC detecting the expression of a tumour associated gene. The present  
 CC sequence is the amino acid sequence of a human PRO protein of the  
 CC invention.  
 XX SQ Sequence 422 AA;

Query Match 36.9%; Score 822; DB 7; Length 422;  
 Best Local Similarity 40.3%; Pred. No. 9.2e-72;  
 Matches 173; Conservative 60; Mismatches 120; Indels 76; Gaps 9;

QY 12 LLDYKTEK-----WALLAKGYQERDLEPQSIITKLG 45  
 DB 13 LFDYKTEKYIAKNKKVGLLYRLLOASTLAYLVVWVFLIKKGYQDVDTLSQSAVITKVG 72  
 QY 46 VSVTQIKELGNRLMDVADFKVPPQGENVFVLTNLTVPQVQGRCPHPSVPLANCWD 105  
 DB 73 VAFNTSDLGRIWDVADYVIPAQGENVFVLTNLTVPQVQGRCPHPSVPLANCWD 132  
 QY 106 EDCPEGEGTHSHGVKTCQCVFNGTHR-TCEIWSWCPVSGVPSRPLLAQAQNTFLFI 164  
 DB 133 SDCHAGEAVTAGNGVTKGRCRLRGNLARGTCEIFAWCPLETSSRPEEPFLKEAEDFTIFI 192  
 QY 165 KNTVTFSKNFPSKNALETWDPYFKHCYEPQSPYCPVFRIGDLVAKAGTFEDLALL 224  
 DB 193 KNHIFPKFNFKNVMDVKDRSFLKSCHEGPK-NHYCPIFRGLSIVRWAGSDFOIALR 251  
 QY 225 GGSVGIRVHWCDDLTGDSGCWPHYSF-QLQEK-----SYNFRATHWWEQGVART 276  
 DB 252 GGVIGINTENWCDLKAASECHPSYFSRSLDNKLSVSSGYNFRFARYRDAAGVEPT 311  
 QY 277 LKLYGIRPDILVTQAGKFGILPATVLTGTAAMLGWVTFPCDLLLIVDREAHFYWRT 336  
 DB 312 LMKYGIIRDVMVNGK-----AFFCDLVLYLKREFFYRDK 349  
 QY 337 KYEEAKAPKATANSVWRLEALASQARLAECRLRSSAPA-----PTATAAGSQ 383  
 DB 350 KYEEVRGLDSSQEADE---ASGLGLEQL---TSGPGLLGWPEQQLQEPPEAKRGSSS 404  
 QY 384 TOTPGWPCP 392

DB 405 QKNGSVCP 413

RESULT 13

ID ADP49181

XX ADP49181 standard; protein; 422 AA.

XX AC ADP49181;

XX DT 26-AUG-2004 (first entry)

XX DE Human P2X5 purinergic receptor protein sequence for odour modulation.

XX KW odour sensitivity; P2X purinergic receptor; P2Y purinergic receptor;

XX KW signal transduction pathway; olfactory signalling; micro-array.

XX OS Homo sapiens.

XX PN WO2004047749-A2.

XX PD 10-JUN-2004.

XX PF 21-NOV-2003; 2003WO-US037389.

XX PR 21-NOV-2002; 2002US-0428140P.

XX PA (UTAH ) UNIV UTAH RES FOUND.

XX PI Lucero M, Hegg C;

XX DR WPI; 2004-460642/43.

XX PT Modulating odour sensitivity in a subject, comprises administering a  
 PT composition comprising an agonist or antagonist of P2X or P2Y purinergic  
 PT receptor to the subject.  
 XX PS Disclosure; SEQ ID NO 10; 108pp; English.  
 XX CC The invention relates to a method of modulating (M1) odour sensitivity in  
 CC a subject, by administering a composition which is an agonist or  
 CC antagonist of a P2X or P2Y purinergic receptor. (M1) is useful for  
 CC modulating odour sensitivity in a subject (claimed). The compositions  
 CC used for modulating odour sensitivity in a subject are useful for  
 CC studying the signal transduction pathways related to olfactory signaling.  
 CC The compositions are also useful as reagents in micro-arrays or as  
 CC reagents to probe or analyze existing micro-arrays. This sequence  
 CC corresponds to the human P2X5 protein sequence.

Sequence 422 AA;

Query Match 36.9%; Score 822; DB 8; Length 422;

Best Local Similarity 40.3%; Pred. No. 9.2e-72;

Matches 173; Conservative 60; Mismatches 120; Indels 76; Gaps 9;

QY 12 LLDYKTEK-----WALLAKGYQERDLEPQSIITKLG 45

DB 13 LFDYKTEKYIAKNKKVGLLYRLLOASTLAYLVVWVFLIKKGYQDVDTLSQSAVITKVG 72

QY 46 VSVTQIKELGNRLMDVADFKVPPQGENVFVLTNLTVPQVQGRCPHPSVPLANCWD 105

DB 73 VAFNTSDLGRIWDVADYVIPAQGENVFVLTNLTVPQVQGRCPHPSVPLANCWD 132

QY 106 EDCPEGEGTHSHGVKTCQCVFNGTHR-TCEIWSWCPVSGVPSRPLLAQAQNTFLFI 164

DB 133 SDCHAGEAVTAGNGVTKGRCRLRGNLARGTCEIFAWCPLETSSRPEEPFLKEAEDFTIFI 192

QY 165 KNTVTFSKNFPSKNALETWDPYFKHCYEPQSPYCPVFRIGDLVAKAGTFEDLALL 224

DB 193 KNHIFPKFNFKNVMDVKDRSFLKSCHEGPK-NHYCPIFRGLSIVRWAGSDFOIALR 251

QY 225 GGSVGIRVHWCDDLTGDSGCWPHYSF-QLQEK-----SYNFRATHWWEQGVART 276

DB 252 GGVIGINTENWCDLKAASECHPSYFSRSLDNKLSVSSGYNFRFARYRDAAGVEPT 311

QY 277 LKLYGIRPDILVTQAGKFGILPATVLTGTAAMLGWVTFPCDLLLIVDREAHFYWRT 336

DB 312 LMKYGIIRDVMVNGK-----AFFCDLVLYLKREFFYRDK 349

QY 337 KYEEAKAPKATANSVWRLEALASQARLAECRLRSSAPA-----PTATAAGSQ 383

DB 350 KYEEVRGLDSSQEADE---ASGLGLEQL---TSGPGLLGWPEQQLQEPPEAKRGSSS 404

QY 384 TOTPGWPCP 392

Db 252 GGVIIGINIENWCDLKAASECHPHYSFRLDNKLSKSVSSYNNFRFARYRDAAQVFRFT 311  
 QY 277 LKLYGIRFDILVTQAGKFGLIPTAVTLGTGAAMLGVVTFPCDLLLLYVDREAHFYWRT 336  
 Db 312 LMKAGIRFDVYNGK-----AFFDLVLIYLIKREFYRDK 349  
 QY 337 KYEERAKAPATANSVWRELALASQARLAECURRSAPA-----PTATAAGSQ 383  
 Db 350 KYEEVRGLEDSQEADE---ASGLGLSEQL--TSGFLLGMPQEQLOPFPPEAKRGSS 404  
 QY 384 TQTPGMPCP 392  
 Db 405 QKNGSVCP 413

RESULT 14  
 AAW47067  
 ID AAW47067 standard; protein; 404 AA.  
 AC AAW47067;  
 XX  
 XX 18-MAY-1998 (first entry)  
 DT  
 DE Human brain P2X-2 receptor polypeptide.  
 KW Brain; P2X receptor; treatment; trauma; spinal cord; infection; human;  
 KW inflammation; mood disorder.  
 OS Homo sapiens.  
 XX  
 XX WO9741222-A1.  
 XX  
 XX 06-NOV-1997.  
 XX  
 XX 30-APR-1996; 96WO-GB001034.  
 XX  
 XX 30-APR-1996; 96WO-GB001034.  
 XX (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX McHale MT, Tomlinson WJ, Livingstone CD, Carpenter DJ, Yi L;  
 XX WPI; 1997-549726/50.  
 XX N-PSDB; AAV09308.  
 XX  
 XX DNA encoding human brain P2X receptors - used to develop products for  
 XX treating, e.g. brain or spinal cord traumas, infection, inflammation and  
 XX mood disorders.  
 XX  
 XX Claim 13; Fig 4; 31pp; English.  
 XX  
 XX This is a human brain P2X-2 receptor polypeptide. The P2X receptor splice  
 XX variants (P2X-1, P2X-2 and P2X-3 isoforms) can be recombinantly expressed  
 XX by a host cell genetically engineered with a vector containing the  
 XX encoding nucleic acids. The receptors can be used to screen for their  
 XX antagonists. The products can be used to identify agents which modulate  
 XX the activity of the receptors for use in clinical conditions such as  
 XX brain stroke, brain or spinal cord traumas, infection and inflammation,  
 XX cognitive disorders, epilepsy, affective and mood disorders in general,  
 XX including depression, various movement disorders including Parkinson's  
 XX disease, Huntingtons Chorea and schizophrenia, as well as those  
 XX conditions that are associated with the development of chronic or acute  
 XX forms of pain, or cardiac anoxia. The products can also be used for  
 XX detection, production of antibodies and production of transgenic animals  
 XX as models for mutation and structure/activity relationship evaluations as  
 XX well as in drug screening

Sequence 404 AA;

Query Match 36.7%; Score 818; DB 2; Length 404;  
 Best Local Similarity 46.8%; Pred. NO. 2.1e-71;  
 Matches 155; Conservative 57; Mismatches 107; Indels 12; Gaps 4;

QY 19 KVALLAKGYQERDLEPOFSIITKLKGVSVTQIKELGNRLMDVADFVKPQGENVFLVLT 78  
 Db 61 RWFVMEKGYQETD-SVSSVTTKYKGVAVTNTSKLGRIMDVADYVIPAQEENSLFVNT 119  
 QY 79 NELVTPAQOQRCPCBHPSPVLANCVWDEDCPEGEGGTHSHGVKTQCQCVVFNGTHTTCBIW 138  
 Db 120 NVILTNQOTGLCPILPDATTV-CKSDASCSTAGSAGTHSNGVSTGRCVAFNGSVKTCVA 178  
 QY 139 SWCPVESGV-VPSRPLLAQAQNTLFIKNTVTTFSEKFNFSKNALETWDTFYFKHCRISPO 197  
 Db 179 AWCPEVDDTHVPQAPAFKAAENFTLLVKNINWYKFNFSKENILFNITTTVLKSCIYDAK 238  
 QY 198 FSPYCPVFRIGDLVAKAGGTFFDLALLGSGSVGIRVHWCDDLTGSGCWPHYSFLOE-- 255  
 Db 239 TDFCFIFRLGKIVENAGHSFQDMAVEGIGIMGIQVWDCNLDRAASLCLPRYSFRRLDTR 298  
 QY 256 -----KSYNFTATHHWEQPVFEARTLLKLYGIRFDILVTGQAGKFGLIPTAVTLGTG 308  
 Db 299 DVEHNVSPGYNFRFAKYRDLAGNEORTLIKAYGIRFDIIVFGKAGKFDIITPTMINIGSG 358  
 QY 309 AAWLGVVTFPCDLLLLYVDREAHFYWRTKYE 339  
 Db 359 LALLGNATVLCDIIVLYCMKKRLYIREKKYK 389

RESULT 15  
 AAW47066  
 ID AAW47066 standard; protein; 388 AA.  
 XX  
 XX AAW47066;  
 XX  
 XX 18-MAY-1998 (first entry)  
 DT  
 DE Human brain P2X-1 receptor polypeptide.  
 XX  
 XX Brain; P2X receptor; treatment; trauma; spinal cord; infection; human;  
 XX inflammation; mood disorder.  
 OS Homo sapiens.  
 XX  
 XX WO9741222-A1.  
 XX  
 XX 06-NOV-1997.  
 XX  
 XX 30-APR-1996; 96WO-GB001034.  
 XX  
 XX 30-APR-1996; 96WO-GB001034.  
 XX (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX McHale MT, Tomlinson WJ, Livingstone CD, Carpenter DJ, Yi L;  
 XX WPI; 1997-549726/50.  
 XX N-PSDB; AAV09307.  
 XX  
 XX DNA encoding human brain P2X receptors - used to develop products for  
 XX treating, e.g. brain or spinal cord traumas, infection, inflammation and  
 XX mood disorders.  
 XX  
 XX Claim 13; Fig 3; 31pp; English.  
 XX  
 XX This is a human brain P2X-1 receptor polypeptide. The P2X receptor splice  
 XX variants (P2X-1, P2X-2 and P2X-3 isoforms) can be recombinantly expressed  
 XX by a host cell genetically engineered with a vector containing the  
 XX encoding nucleic acids. The receptors can be used to screen for their  
 XX antagonists. The products can be used to identify agents which modulate  
 XX the activity of the receptors for use in clinical conditions such as  
 XX brain stroke, brain or spinal cord traumas, infection and inflammation,  
 XX cognitive disorders, epilepsy, affective and mood disorders in general,  
 XX including depression, various movement disorders including Parkinson's  
 XX disease, Huntingtons Chorea and schizophrenia, as well as those







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 8628.1 Seconds  
(without alignments)  
11373.559 Million cell updates/sec

Title: US-09-820-095B-1  
Perfect score: 2693  
Sequence: 1 ttgtgactcatgtgcgcg.....aaaaaaaaaaaaaaaaaaaaa 2693

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hsc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1508	56.0	2191	3	BC047287 Homo sapi
2	857.8	31.9	1068	5	BX338022 BX338022
3	772.8	28.7	2222	3	AK081308 Mus muscu
4	751	27.9	787	4	BG920544 602782887
5	740.8	27.5	850	8	AQ748500 HS-5540_A
6	737.4	27.4	1022	5	BX359188 BX359188
7	640.6	23.8	682	2	BF529980 602042384
8	584.2	21.7	3383	3	AK054195 Mus muscu
9	576.2	21.4	741	4	BI821165 603035061
10	561	20.8	733	4	BI753477 603026391
11	553	20.5	718	6	CF131935 UI-HF-PQ0
12	552	20.5	552	4	BM96678 UI-E-DW0
13	472.8	17.6	1038	4	BI757011 603028734
14	466.4	17.3	477	2	AW293881 UI-H-BW0
15	450.4	16.7	460	4	BM689096 UI-E-CQ1
16	445.4	16.5	458	1	AI096436 qbs9d11.x
17	417.2	15.5	422	4	BM668009 UI-E-DW0
18	371.4	13.8	651	5	BO210100 UI-R-EP0
19	361.4	13.4	637	6	BY728336 BY728336
20	359.4	13.3	612	6	BY724280 BY724280
21	349.2	13.0	357	1	AJ573523 AJ573523
22	347	12.9	497	8	AQ266759 RPC111-76
23	338	12.6	1273	5	BO706537 AGENCOURT
24	337.8	12.5	1071	5	BQ894752 AGENCOURT

## ALIGNMENTS

RESULT 1  
BC047287  
LOCUS BC047287 2191 bp mRNA linear HTC 03-MAR-2003  
DEFINITION Homo sapiens, purinergic receptor P2X-like 1, orphan receptor,  
clone IMAGE:5198937, mRNA.

ACCESSION BC047287 GI:28703829

VERSION BC047287.1

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2191)

AUTHORS Strausberg R.

TITLE Direct Submission

JOURNAL Submitted (28-FEB-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT Contact: MGC help desk

Email: gcgaps-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: http://www.nisc.nih.gov/

Contact: nisc\_mgc@nigri.nih.gov

Akhtar, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masillo, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McGowan, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Teague, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAC Plate: 93 Row: i Column: 7

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 4885534

This clone has the following problem: frame shifted.

FEATURES Location/Qualifiers

1. .2191

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5198937"  
 /tissue\_type="Brain, adult, 6 pooled whole brains"  
 /clone\_lib="NIH MGC\_114"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

## ORIGIN

Query Match 56.0%; Score 1508; DB 3; Length 2191;  
 Best Local Similarity 98.7%; Pred. No. 0;  
 Matches 1520; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 62 ACAGGCTGGGGCTCTCGATTATAGACGAGAGTGGGCTCTCTCGCCAAAAGGC 121  
 DB 113 AGAGCTGCTGAGTTGGGATCGTGCTATGTGGGCTCTCTCGCCAAAAGGC 172  
 QY 122 TACAGGACGGGACCTGGAAACCCAGTTTTCATCATACCAAACTCAAAGGGGTTTC 181  
 DB 173 TACCAGGACGGGACCTGGAAACCCAGTTTTCATCATACCAAACTCAAAGGGGTTTC 232  
 QY 182 GTCACTCAGATCAAGGAGCTTGGAAACCGGCTGGGATGGCCGACTTCGTGAAGCCA 241  
 DB 233 GTCACTCAGATCAAGGAGCTTGGAAACCGGCTGGGATGGCCGACTTCGTGAAGCCA 292  
 QY 242 CCTCAGGAGAGAAAGTGTCTTCTTGGTGACCAACTTCCCTGTGACGCCAGCCCAAGTT 301  
 DB 293 CCTCAGGAGAGAAAGTGTCTTCTTGGTGACCAACTTCCCTGTGACGCCAGCCCAAGTT 352  
 QY 302 CAGGACATCCAGAGACCCCTCGTCCCACTGCTAACTGCTGGGTGACGAGGAC 361  
 DB 353 CAGGACATCCAGAGACCCCTCGTCCCACTGCTAACTGCTGGGTGACGAGGAC 412  
 QY 362 TGCCCCGAAGGGGAGGACACACAGCCACCGGTGTAAACACAGCCCAAGTGTGGTG 421  
 DB 413 TGCCCCGAAGGGGAGGACACACAGCCACCGGTGTAAACACAGCCCAAGTGTGGTG 472  
 QY 422 TTCAATGGACCCAGAGACCTGTGAGATCTGGAGTTGGTCCAGTGGAGAGTGGGTT 481  
 DB 473 TTCAATGGACCCAGAGACCTGTGAGATCTGGAGTTGGTCCAGTGGAGAGTGGGTT 532  
 QY 482 GTGCCCTCAGGCCCCCTGCTGGCCAGCCAGACCTTCACTGTTTCATCAAAAACACA 541  
 DB 533 GTGCCCTCAGGCCCCCTGCTGGCCAGCCAGACCTTCACTGTTTCATCAAAAACACA 592  
 QY 542 GTCACTTCAGCAAGTTCACATTTCTAAGTCCAAATGCTTGGAGACCTGGGACCCACC 601  
 DB 593 GTCACTTCAGCAAGTTCACATTTCTAAGTCCAAATGCTTGGAGACCTGGGACCCACC 652  
 QY 602 TATTTTAAGCACTCCGCTATGACCAATTCAGCCCTACTGTCCTGGTGTCCGCAAT 661  
 DB 653 TATTTTAAGCACTCCGCTATGACCAATTCAGCCCTACTGTCCTGGTGTCCGCAAT 712  
 QY 662 GGGACCTCTGTGGCAAGGCTGGAGGACCTTCAGGACCTGGCGTTGTCTGGTGGCTCT 721  
 DB 713 GGGACCTCTGTGGCAAGGCTGGAGGACCTTCAGGACCTGGCGTTGTCTGGTGGCTCT 772  
 QY 722 GTAGGCATCAGATTCACCTGGGATGTGACCTGGACACCGGGACCTGGCTGTGGCT 781  
 DB 773 GTAGGCATCAGATTCACCTGGGATGTGACCTGGACACCGGGACCTGGCTGTGGCT 832  
 QY 782 CACTACTCTCTTCCAGCTGCAGGAGAGAGTCAAACTTCAGGACAGCCACTCACTGGTG 841  
 DB 833 CACTACTCTCTTCCAGCTGCAGGAGAGAGTCAAACTTCAGGACAGCCACTCACTGGTG 892  
 QY 842 GAGCAACGGGTGTGGAGGCCCGCACCTCTGCTAAGCTCTATGAATCCGCTTCGACATC 901  
 DB 893 GAGCAACGGGTGTGGAGGCCCGCACCTCTGCTAAGCTCTATGAATCCGCTTCGACATC 952  
 QY 902 CTCGTCACCGGGCAGGAGGAAAGTTCGGGCTCATCCCAACCGGCTCACTGGGACCC 961  
 DB 953 CTCGTCACCGGGCAGGAGGAAAGTTCGGGCTCATCCCAACCGGCTCACTGGGACCC 1012

QY 962 GGGCAGCTTGGCTGGGGTGGTCACTTTTCTGTGACCTGCTACTGCTGTATGTGGAT 1021  
 DB 1013 GGGCAGCTTGGCTGGGGTGGTCACTTTTCTGTGACCTGCTACTGCTGTATGTGGAT 1072  
 QY 1022 AGAAGAGCCCATTTTCTACTGGAGGACAAAGTATAGAGGCGCAAGGCCCGGAAAGCAACC 1081  
 DB 1073 AGAAGAGCCCATTTTCTACTGGAGGACAAAGTATAGAGGCGCAAGGCCCGGAAAGCAACC 1132  
 QY 1082 SCCAATCTGTGTGGAGGAGCTGGCCCTTGCATCCCAAGCCGACTGGCGAGTGGCTC 1141  
 DB 1133 GCCAATCTGTGTGGAGGAGCTGGCCCTTGCATCCCAAGCCGACTGGCGAGTGGCTC 1192  
 QY 1142 AGACGGAGCTCAGCAGCTGCACCCACGGCCACTGTCTGTGGAGTCAAGACACACCA 1201  
 DB 1193 AGACGGAGCTCAGCAGCTGCACCCACGGCCACTGTCTGTGGAGTCAAGACACACCA 1252  
 QY 1202 GGATGGCCCTGTCCAAGTCTGACACCCACTTGCACACCCATTCGGGAGCTGTAGCG 1261  
 DB 1253 GGATGGCCCTGTCCAAGTCTGACACCCACTTGCACACCCATTCGGGAGCTGTAGCG 1312  
 QY 1262 TTCCCTGTGTGTGAGAGTTGGGGCTGGGAAGGGCGGGCCCTGCTGGGGATCTCAAG 1321  
 DB 1313 TTCCCTGTGTGTGAGAGTTGGGGCTGGGAAGGGCGGGCCCTGCTGGGGATCTCAAG 1372  
 QY 1322 GATGAGGCCCGAGCATGGAGGATTTGGGGTGAATTCACCCCTTGAACCCCGAGACAG 1381  
 DB 1373 GATGAGGCCCGAGCATGGAGGATTTGGGGTGAATTCACCCCTTGAACCCCGAGACAG 1432  
 QY 1382 TCCTTCCTCTGACTCCCACTTGGTGGTGTGCTGCTCAGGAGCCATAGAAGTGGGCTG 1441  
 DB 1433 TCCTTCCTCTGACTCCCACTTGGTGGTGTGCTGCTCAGGAGCCATAGAAGTGGGCTG 1492  
 QY 1442 TGTTTTGAAGCGGACAGAACTGACCCGCTGGAGACTGGGAGAGCCCGAGGACCTG 1501  
 DB 1493 TGTTTTGAAGCGGACAGAACTGACCCGCTGGAGACTGGGAGAGCCCGAGGACCTG 1552  
 QY 1502 TATTCAGGCTCCGACTGTCATGTGGAGGGCTCTGCTCGCTGGGCTGGAGGCTCT 1561  
 DB 1553 TATTCAGGCTCCGACTGTCATGTGGAGGGCTCTGCTCGCTGGGCTGGAGGCTCT 1612  
 QY 1562 CTCTCCCTGCTGCTGTCTCCCACTGTTCTTACGAGAGGTA 1601  
 DB 1613 CTCTCCCTGCTGCTGTCTCCCACTGTTCTTACGAGAGGAA 1652

RESULT 2  
BX338022/c

LOCUS BX338022 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS01054YAL2 3-PRIME, mRNA sequence.

## DEFINITION

BX338022

## ACCESSION

BX338022

## VERSION

BX338022.2

## KEYWORDS

EST.

## ORGANISM

Homo sapiens (human)

## REFERENCE

1 (bases 1 to 1658)

## AUTHORS

Li, W.-B., Gruber, C., Jesse, J., and Polayes, D.

## TITLE

Full-length cDNA libraries and normalization

## JOURNAL

Unpublished (2001)

## COMMENT

On May 2, 2003 this sequence version replaced gi:30346991.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

10471.f

```

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0D1054BA06NP1&c=10471.f.
Location/Qualifiers
1. .1068
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1054YA12"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	31.9%	Score 857.8;	DB 5;	Length 1068;
Best Local Similarity	94.6%;	Pred. No. 9.8e-195;		
Matches 919;	Conservative 15;	Mismatches 32;	Indels 5;	Gaps 5;
Qy	1650	CTCTCCCCCATCTGCAACCCCACTATAGGTAGAGACCCACCCCTCCATCGGTCTTACA	1709	
Db	969	CCCCCCCCCAYCGCMCCCCCAITCAAAGG-AGARCCCCCACCCCCCBGCGGCCDACA	911	
Qy	1710	TGGGGCTGTCGACTGGAGCCAAAAGCGCAGGCGAGAGAGAGTGTGCGGAGGAGGGG	1769	
Db	910	-GGGGGGGGGGGGGAGCTTTTWKCAAGGTAGAAAGGAGTGTATGGGGAGGGGG	852	
Qy	1770	ATTGTTTCAGCTTCTCTGTGCTGTGATGCCCCAGGAGGTCTTAATCTAGGGAATGGG	1829	
Db	851	ATTGTTTCAGCTTCTCTGTGCTGTGATGCCCCAGGAGGTCTCTTAATCTAGGGAATGGG	792	
Qy	1830	TGGAGTAGGCAGATAAATCCACCTCCCTATCCCCACAGCAAGGGCGGAGCATGTCTCTGG	1889	
Db	791	TGGAGTAGGCAGATAAATCCACCTCCCTATCCCCAGCAGAGGGCGGAGCATGTCTCTGG	732	
Qy	1890	GCCACACCTGCTTAGTTTATGAGACCGGTGTCTTCCAGTGTAGCCCTTTTGGCATG	1949	
Db	731	GCCACACCTGCTTAGTTTATGAGACCGGTGTCTTCCAGTGTAGCCCTTTTGGCATG	672	
Qy	1950	GAGTCTGGGAGACAGACGAGCGGCGGCGAGGGCTAAGTTG-GTGATCATTTGGGTTCTTC	2008	
Db	671	GAGTCTGGGAGACAGACGAGCGGCGGCGAGGGTAAAGTTGTGTGATCATCTGGGTTCTTC	612	
Qy	2009	AGGACCTTCTATATCCCTCTCTCGGTAAACCCCCACGCCCAACCCCTTGGAAATCTTCTCTCC	2068	
Db	611	AGGACCTTCTATATCCCTCTCTCGGTGACCCCCACGCCCAACCCCTTGGAAATCTTCTCTCC	552	
Qy	2069	AGGCTTCCTGAGAGCCCTGGGGGTGGAGGCTGTGGAGGCTGTACATCTGAAATTCATCT	2128	
Db	551	AGGCTTCCTGAGAGCCCTGGGGGTGGAGGCTGTGGAGGCTGTACATCTGAAATTCATCT	492	
Qy	2129	TCAGTCCAAAGTCATACCTAGGAAGCTGTCTGGGCAGCTGTCTCGAGGAGGCCCTTGGCTCT	2188	
Db	491	TCAGTCCAAAGTCATACCTAGGAAGCTGTCTGGGCAGCTGTCTCGAGGAGGCCCTTGGCTCT	432	
Qy	2189	GATCCCAAGCTGGATGGAGTGGCTGGAAGGAATGGTCCAAACCAACACCCAGAGATCTC	2248	
Db	431	GATCCCAAGCTGGATGGAGTGGCTGGAAGGAATGGTCCAAACCAACACCCAGAGATCTC	372	
Qy	2249	CCTCAGGCTGGCCAGGTTTTTCAGCTGTGAAATTTCTCTCTTTGGTCCAGGGCGGGCGAGGG	2308	
Db	371	CCTCAGGCTGGCCAGGTTTTTCAGCTGTGAAATTTCTCTCTTTGGTCCAGGGCGGGCGAGGG	312	
Qy	2309	AAATCTAAGTGTCCACCCCAAGGAGGCAAGGGGCTGCTTTCCACTGTGGGTAACCTGGTGA	2368	
Db	311	AAATCTAAGTGTCCACCCCAAGGAGGCAAGGGGCTGCTTTCCACTGTGGGTAACCTGGTGA	252	
Qy	2369	TCAGGCAAGCTGTGGAGGGCCAGGGTGGGGCTGTAGACTGGGCTGACATCTAGAAATCAC	2428	
Db	251	TCAGGCAAGCTGTGGAGGGCCAGGGTGGGGCTGTAGACTGGGCTGACATCTAGAAATCAC	192	
Qy	2429	CTGCAACTGTGAGGCTCTAGTAAATGCTGTGGGGTCCCTGTCTGCTCTCAATCTCCAGAGC	2488	

[illegible][illegible]

4 THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
REFERENCE  
5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
JOURNAL Nature 420, 563-573 (2002)  
REFERENCE  
6 (bases 1 to 2222)  
AUTHORS Fuchida,S., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Adachi,J., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Oheato, N., Okazaki, Y., Saio, R., Saich, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

## TITLE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

URL: <http://fantom.gsc.riken.jp/>

## FEATURES

source  
Location/Qualifiers  
1. .2222  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM,DB:C130003C22"  
/db\_xref="taxon:10090"  
/clone="C130003C22"  
/tissue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="16 days embryo"  
36. .1176  
/note="purinergic receptor P2X-like 1, orphan receptor (MGD|MG1:1337113, GB|NM\_011028, evidence: BLASTN, 100%, match=801)  
putative"

## misc\_feature

Query Match 28.7%; Score 772.8; DB 3; Length 2222;  
Best Local Similarity 75.9%; Pred. No. 3e-174;  
Matches 1011; Conservative 0; Mismatches 307; Indels 14; Gaps 4;

## ORIGIN

97 GTGGGCTCTCTCCGCAAAAAGGCTACAGAGCGGACCTGGAAACCCAGTTTCCAT 156  
176 GTGGGCTCTCTGGCCCAAAAAGGCTACAGAGCGGACCTGGCCCTCAGACTTCTGT 235  
157 CATCAAACTCAAGGGTTTCCGCTCACTCAGATCAAGAGCTTGGAAACCGGTGTG 216  
236 CATCAAACTCAAGGGTTTCTGTAACTCAGGTTAAGAGCTGAGAACCGGTGTG 295  
217 GGATGTGGCGACTTCGTGAAGCCACCTCAGGGAGAGACGTGTTCTTCTGTGACCAA 276  
296 GGAGCTGGCTGACTTGTGAAGCCATCACAGGAGAGACGTGTTCTTCTGTTAACCAA 355  
277 CTTCTTGTGACCGCCAGCTTCTAGGAGAGTCCAGAGACACCGCTCCCTCCACT 336  
356 CTTCTTGTGACACCACTCAAGTCCAGGAGAGTCCAGAGACATCTTCCCTTCTCT 415  
337 GGCTAACTGTGGTTCGACGAGGACTCCCGGAGGGGAGGAGGACACACAGCCACCG 396  
416 GGCTAACTGTGGCTGATGAGACTGCCCGAGGGGAGACCGGGACATACAGCCATG 475  
397 TGTAAAAACAGCCAGTGTGTGTGTTCAATGGAGCCACAGGACCTGTGAGATCTGGAG 456  
476 CATAAAAAAGTGTGTGTGTGTGTTCAACGGAACCCACAGGACCTGTGAGATCTGGAG 535  
457 TTGTGCGCCAGTGGAGAGTGGCGTTGTCCTCGAGGCCCTCTGAGCCCGCCAGAA 516  
536 CTGGTGGCCAGTGGAGAGTGGTCTGTGCGCCAGGAGCCCTCTGCTAGCTCAGGCCAAGAA 595

517 CTTCACTGTCTCATCAAAAACACAGTCACTCTCAGCAAGTTCAACTTCTCTAAGTCCAA 576  
596 CTTCACTGTCTCATCAAAAACACAGTCACTCTCAGCAAGTTCAACTTCTCTAAGTCCAA 655  
577 TGCCTTGGAGAGCTGGGACCCACCTATTTAAGCACTGCGGCTATGAACACCAATTCAG 636  
656 TGCCTTATTAACCTGGGACACACATATTTCAAGCAATTTCTCTATGATCCACTCTCAG 715  
637 CCCCTACTGTCCCGGTTCCGATTTGGGACCTCTGTCGCAAGGCTGGAGAGACCTTCGA 696  
716 TCCGTACTGCCAGTGTTCGCACTCGGACCTGTGTGCCATGGCTGGAGGGGACTTGA 775  
697 GGACCTGGCGTT-GCTGGGTGGCTCTGTAGGCATCAGAGTTCACTGGGATTTGACCTGG 755  
776 GGACCTGGCGTTGGCTGGGTGGTGTGGGCATCAGCATTTCACTGGGATTTGCAACCTGG 835  
756 ACACGGGGACTTGGCTGGCTGGCTCACTACTCTTCAGCTCAGAGAGAGAGTACA 815  
836 ACACGAAAGGCTTCACTGTCTGCCCCAGTACTCTTCCAGTGGCAGAGAGAGTACA 895  
816 ACTTCAGACAGCCACTCACTGGTGGGAGCAACCGGTTGTGGAGCCCGCACCTGTCTCA 875  
896 ACTTCAGACAGCCACTCACTGGTGGGAGCCCTCAGGCTGGAGACCGGAGCTGCTCA 955  
876 AGCTCTATGGAAATCCGCTTCGACATCTCTGTCACGGGACAGGAGGAGTTCGGGCTCA 935  
956 AGCTCTATGGAAATCCGCTTTTGACATTTCTGTCACTGGGACAGGAGGAGTTCGGCTCA 1015  
936 TCCCAACGGCGTCACTGGGACCGGGGAGCTTGGCTGGGCTGTGCTACCTTTTCT 995  
1016 TCCCTACGGCCATCACTGGGACAGGAGGCTTGGCTGGGATGGTCACTTCTCT 1075  
996 GTGACCTGCTACTGTGTATGTGATAGAGAGCCATTTCTACTGGAGGACAAAGTATG 1055  
1076 GTGACCTGCTACTGTATGTGATAGAGAGCGGTTTCTACTGGAGAGGAGTTCGCTTGCAT 1135  
1056 AGGAGCCAAAGCCCGGAAAGCAACCGGCTCTGTGTGGAGGAGTGGCCCTTGCAT 1115  
1136 AGAGGCAAGAGCCCAAGAGAGCTACCAAGATTTCTACTGGAGAGGAGGAGTTCAC 1195  
1116 CCCAAGCCGAGTGGCCGAGTGCCTCAGACGAGCTCAGACCTTCAGACCCAGGAGTGG 1175  
1196 CCGCTAACCGTGTACCCAGTGTCTTAGGAGTGGTTTGACCTGC-CCCATGTCTGTG 1254  
1176 GTGCTGGAGTCAAGACAGA-----CACCAGATGGCCCTGTCCAAGTCTTGAC 1225  
1255 GTGCTGGGATCCAGGCAATCCAGCCAGCCAGCCAGGATGGCTGTCTGGATGGCTGCC 1314  
1226 ACCCACTTGCACAAACCCATTCCGGGAGCTGTAGCGCTTCCCTGTGTGGTGGAGTGGGG 1285  
1315 ACCCACTGGCTATTCATTTCTGAGGCTTACACCTCTCCCTGTGTGTTAAAGTTAGGG 1374  
1286 GCTGGGAAGG--GGGGGCCCTGCTGGGATCTCAAGGATGAGGCCCGGAGTGGAGGA 1343  
1375 CTTGGGAAGGAGGGTTCGGCTTGGGAAATGGAGGACAAAGCTTCAGCGGAGTAG 1434  
1344 TTGGGGGTAGAAATCCACCTTGAACCCAGACAGTCCCTCCCTGACTGCCACCTT 1403  
1435 AGGGAGTTGGATTAATTCGACAGACTCCGANTGCAAGATTCCTCTCTGGTTTCACTAG 1494  
1404 GGTAGGGTGTG 1415  
1495 AGGGTGGGGTG 1506

## RESULT 4

BG820544  
LOCUS 60278287F1 NCI\_CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4933821  
DEFINITION 787 bp mRNA linear EST 22-MAY-2001  
5', mRNA sequence.  
ACCESSION BG820544  
VERSION BG820544.1 GI:14168131  
KEYWORDS EST.

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgaabs@mail.nih.gov
              Tissue Procurement: David N. Louis, M.D.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (BLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/BLNL at:
              http://image.llnl.gov
              Plate: LIAW10861 row: m column: 22
              High quality sequence stop: 733.

FEATURES    source
            1..787
            /organism="Homo sapiens"
            /mol_type="rRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4933821"
            /tissue_type="anaplastic oligodendroglioma with lp/19q
            loss"
            /lab_host="DH10B (T1 phage-resistant)"
            /clone_lib="NCI CGAP Brn67"
            /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
            Average insert size 2.3 kb. Constructed by Life
            Technologies. Note: this is a NCI CGAP Library."

ORIGIN
Query Match 27.9%; Score 751; DB 4; Length 787;
Best Local Similarity 98.7%; Pred. No. 4.2e-169;
Matches 757; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1913 GCACCGGTGCTTCCAGTGTAGCCCTTTTGCATGGAGGTCTGGGAGAGACAG 1972
DB 1 GCACCGGTGCTTCCAGTGTAGCCCTTTTGCATGGAGGTCTGGGAGAGACAG 60

QY 1973 GCGGCGAGGCTAAGTGGTGATCTGGTTCCTTCAGACCTTCATATCCCTCTCG 2032
DB 61 GCGGCGAGGCTAAGTGGTGATCTGGTTCCTTCAGACCTTCATATCCCTCTCG 120

QY 2033 TAAACCCCGAGCCCAACCCCTTGGAAATCTTCCCTCCAGGCTTCTCGAGCCCTGGGGGT 2092
DB 121 TAAACCCCGAGCCCAACCCCTTGGAAATCTTCCCTCCAGGCTTCTCGAGCCCTGGGGGT 180

QY 2093 GGGAGGCTGTGGAGGCTGTACATCTGAATTCATTGACCTCAAGTCAATACCTAGGAAG 2152
DB 181 GGGAGGCTGTGGAGGCTGTACATCTGAATTCATTGACCTCAAGTCAATACCTAGGAAG 240

QY 2153 CTGCTGGGAGCTCTCGAGGGAGCCCTGGCTCTGATCCAGGCTGGATGGAGTGGCT 2212
DB 241 CTGCTGGGAGCTCTCGAGGGAGCCCTGGCTCTGATCCAGGCTGGATGGAGTGGCT 300

QY 2213 GGAAGGAATGTTCCAAACCAACACCGAGATCTCCCTCAGGCTGGCCAGGTTTTCAG 2272
DB 301 GGAAGGAATGTTCCAAACCAACACCGAGATCTCCCTCAGGCTGGCCAGGTTTTCAG 360

QY 2273 CTGGAATTCCTCTTGTGTCAGGCGGGGAGGAAATTCATAGTTCACCCAGGGA 2332
DB 361 CTGGAATTCCTCTTGTGTCAGGCGGGGAGGAAATTCATAGTTCACCCAGGGA 420

QY 2333 GCGAAGGGGCTGCTTTCACCTGTGGTATCCAGTGGTATCAGGCGGAGCTGTGGAGGGCCAG 2392
DB 421 GCGAAGGGGCTGCTTTCACCTGTGGTATCCAGTGGTATCAGGCGGAGCTGTGGAGGGCCAG 480

QY 2393 GGGTGGGCTGAGACTGGGCTGACATCTAGAATCACTGCCACTCGAGGCTCAGTAAAA 2452

```

```

DB 481 GGGTGGGCTGAGACTGGGCTGACATCTAGAATCACCTGCCACTGAGGCTCAGTAAAA 540
QY 2453 TGCCTGGGGTCCCGCTGCGCTTCAATCTCCAGAGCCATGTCATGGGGAGGTGGGTCT 2512
DB 541 TGCCTGGGGTCCCGCTGCGCTTCAATCTCCAGAGCCATGTCATGGGGAGGTGGGTCT 600
QY 2513 GAAGGGGCAAGTGGGAGAGCAGGCGCCCTGAGGCTGGGTATCCAGAGGGGCGACGTG 2572
DB 601 GAAGGGGCAAGTGGGAGAGCAGGCGCCCTGAGGCTGGGTATCCAGAGGGGCGACGTG 660
QY 2573 CACCTGATTCCTTGGGGCCAGAGGAGCTGATGTCATGGGTGGGACAAAGTCACGGAG 2632
DB 661 CACCTGATTCCTTGGGGCCAGAGGAGCTGATGTCATGGGTGGGACAAAGTCACGGAG 720
QY 2633 TAAAGCCAGCAAGCCACCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2679
DB 721 TAAAGCCAGCAAGGTCACCTTAAAAAACAACCAAGGAGGAGGAGGAGGAGGAGGAG 767

RESULT 5
LOCUS      AQ748500/c
DEFINITION HS_5540_A2_G07_T7A_RPCI-11 Human Male BAC Library Homo sapiens
ACCESSION  AQ748500
VERSION     AQ748500.1
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    1 (bases 1 to 850)
AUTHORS      Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
              Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
              Hood,L.
              Sequence-tagged connectors: A sequence approach to mapping and
              scanning the human genome
              Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
              93380589
              PUBMED 1049764
              Contact: Mahairas GG, Wallace JC, Hood L
              High Throughput Sequencing Center
              University of Washington
              401 Queen Anne Avenue North, Seattle, WA 98109, USA
              Tel: (206) 616-3618
              Fax: (206) 616-3887
              Email: jwallace@u.washington.edu
              Clones are derived from the human BAC library RPCI-11. For BAC
              library availability, please contact Pieter de Jong
              (pieterdejong.med.buffalo.edu). Clones may be purchased from
              BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)
              or from Research Genetics (info@resgen.com). BAC end Web Server:
              http://www.htsc.washington.edu
              Plate: 1116 row: M column: 14
              Seq primer: T7
              Class: BAC ends
              High quality sequence stop: 850.
              Location/Qualifiers
                1..850
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /clone="Plate=1116 Col=14 Row=M"
                /sex="male"
                /notes="RPCI-11 Human Male BAC Library"
                /vector="pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
                Male blood DNA was isolated from one randomly chosen donor
                and partially digested with a combination of EcoRI and
                EcoRI Methylase. Size selected DNA was cloned into the
                pBACe3.6 vector at EcoRI sites"

ORIGIN
Query Match 27.5%; Score 740.8; DB 8; Length 850;

```









Qy	1117	CCAAAGCCGGACTGGCCGAGTGCCTCAGACGGAGCTCAGCACTGCACCCACGGCCACTGC	1176
Db	1163	CGCTAAACCCCTGTATCCCAAGTGTCTTAGGAGTGGTTTGCACTGC-CCCATGCTGTGCTGC	1221
Qy	1177	TGCTGGGAGTCAGACACAGA-----CACCAGGATGGCCCTGTCCTCAAGTTCAGACA	1226
Db	1222	TGCTGGGATCCAGGCATCCAGCACGACCGACCGATGGCTGTCTGATGGCTGCCA	1281
Qy	1227	CCCACTTGGCCAAACCCATTCGGGAGCCTGTAGCGCTTCCCTGCTGTTGAGAGTTGGGGG	1286
Db	1282	CCCACTGGCCTATTTCATCTCGAGGCTTACACCTCTCCCTGCTGGTAAAGTTTAGGGC	1341
Qy	1287	CTGGGAGG--GCGGGCCCTGCTGGGGATCTCAGGATGAGGCCCCACCATGAGGAT	1344
Db	1342	TTGGGAAGGAAGAGGTTCTGGCTTGGGAAAAATGGAGACAAGACTTCACCGGGACTAGA	1401
Qy	1345	TGGGGGTAGAATTCCACCCTTGAACCCAGCAGACAGTCCCTCCCTGTACTCCCACTTG	1404
Db	1402	GGGAGTGTGGATAATTCGACAGACTCCGAATGCAGATTCCCTCTGGGTTTCACTCAGA	1461
Qy	1405	GTAGGGTGTCTG	1415
Db	1462	GGGGTGGGGTG	1472

[illegible]

```

FEATURES
  source
    .741
    Location/Qualifiers
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5176046"
      /lab_host="DH10B"
      /clone_lib="NIH MGC 115"
      /note="Organ: Pooled brain, lung, testis; Vector:
      pcwv-sPORT6; Site 1: NotI, site 2: EcoRV (destroyed); RNA
      source anonymous pool of 6 male brains, age range 23-27, 1
      male lung, age 27; and 1 male testis, age 69. Library is
      oligo-dT primed and directionally cloned (EcoRV site is
      destroyed upon cloning). Average insert size 1.8 kb,
      insert size range 1-3 kb. Library is normalized and
      enriched for full-length clones and was constructed by C.
      Gruber (Invitrogen). Research Genetics tracking code
      021. Note: this is a NIH MGC Library."

```

## ORIGIN

Query Match	21.4%;	Score 576.2;	DB 4;	Length 741;
Best Local Similarity	99.3%;	Prod. No. 4e-127;		
Matches 589;	Conservative 0;	Mismatches 3;	Indels 1;	Gaps 1;
Qy	97	GTGGGCTCTCTCGCCCAAAAAGGCTACCAAGAGCGGACCTGGAAACCCAGTTTTCAT	156	
Db	149	GTGGGCTCTCTCGCCCAAAAAGGCTACCAAGAGCGGACCTGGAAACCCAGTTTTCAT	208	
Qy	157	CATCACCAAACTCAAAAGGGTTTCGGTCACTCAGATCAAGAGCTTGGAAACCGGCTGTG	216	
Db	209	CATCACCAAACTCAAAAGGGTTTCGGTCACTCAGATCAAGAGCTTGGAAACCGGCTGTG	266	
Qy	217	GGATGTGGCCGACTTCGTGAAGCCACCTCAGGGAGAGAACGTGTCTCTTGTGTGACCAA	276	
Db	269	GGATGTGGCCGACTTCGTGAAGCCACCTCAGGGAGAGAACGTGTCTCTTGTGTGACCAA	328	
Qy	277	CTTCCTTTGTGACGCCACGCCAAGTTTCAGGGCAGATGCCCAGACACCGTCCGTCCTCACT	336	
Db	329	CTTCCTTTGTGACGCCACGCCAAGTTTCAGGGCAGATGCCCAGACACCGTCCGTCCTCACT	388	
Qy	337	GGCTAACTGTCTGGTGTGACGAGGACTGCCCGAAGGGGAGGGAGGCAACACAGCCACGG	396	
Db	389	GGCTAACTGTCTGGTGTGACGAGGACTGCCCGAAGGGGAGGGAGGCAACACAGCCACGG	448	
Qy	397	TGTAAAAACAGGCCAGTGTGTGGTGTCAATTTGGGACCCACAGGACCTGTGAGATCTGGAG	456	
Db	449	TGTAAAAACAGGCCAGTGTGTGGTGTCAATTTGGGACCCACAGGACCTGTGAGATCTGGAG	508	
Qy	457	TTGGTGCCCACTGGAGAGTGGGCTTGTGCCCTCGAGGCCCTTGC - TGGCCACAGGCCGAGA	515	
Db	509	TTGGTGCCCACTGGAGAGTGGGCTTGTGCCCTCGAGGCCCTTGC - TGGCCACAGGCCGAGA	568	
Qy	516	ACTTCACACTGTTTCATCAAAAAACACAGTCACCTTCAGCAAGTTCACACTTCTCTAAGTCCA	575	
Db	569	ACTTCACACTGTTTCATCAAAAAACACAGTCACCTTCAGCAAGTTCACACTTCTCTAAGTCCA	628	
Qy	576	ATGCCITGGAGACTGGGACCCCACTATTATTTAAGCACTGGCTATGACCCAAATTCAC	635	
Db	629	ATGCCITGGAGACTGGGACCCCACTATTATTTAAGCACTGGCTATGACCCAAATTCAC	688	
Qy	636	GCCTCTACTCTCCGTTTCCTCGATTGGGACCTCTGTGGCCAAAGGCTGGAGGG	688	
Db	689	GCCTCTACTCTCCGTTTCCTCGATTGGGACCTCTGTGGCCAAAGGCTGGAGGG	741	

RESULT 10  
 BI753477  
 LOCUS  
 DEFINITION  
 733 bp mRNA linear EST 25-SEP-2001  
 603026391F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5196695 5',  
 mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 EST.  
 BI753477.1 GI:15745055  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 733)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapb@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM11492 row: f column: 24  
High quality sequence stop: 719.  
Location/Qualifiers

FEATURES

## FEATURES

```

source
1. 733
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5196695"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/site="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 20.8%; Score 561; DB 4; Length 733;
Best Local Similarity 82.9%; Pred. No. 1.8e-123;
Matches 719; Conservative 0; Mismatches 0; Indels 148; Gaps 1;

QY 2 TGCTGACTCATGTCGCCGACAGCTAGCAGGAGCTGGCAGCATGGCTCCCGAGGGCTAGC 61
DB 14 TGCTGACTCATGTCGCCGACAGCTAGCAGGAGCTGGCAGCATGGCTCCCGAGGGCTAGC 73

QY 62 ACAGCTGGGGCTCTGGATTATAAGACGAGAGTGGGCTCTCTCTCGCCAAAGAGGC 121
DB 74 ACAGCTGGGGCTCTGGATTATAAGACGAGAGTGGGCTCTCTCTCGCCAAAGAGGC 133

QY 122 TACCAGAGCGGACCTGGAAACCCAGTTTCATCATACCAAACTCAAGGGGTTTCC 181
DB 134 TACCAGAGCGGACCTGGAAACCCAGTTTCATCATACCAAACTCAAGGGGTTTCC 193

QY 182 GTCACTCAGTCAAGGAGCTGGAAACCCGCTGGGATGTGGCGGACTTCTGTAAGCCA 241
DB 194 GTCACTCAGTCAAGGAGCTGGAAACCCGCTGGGATGTGGCGGACTTCTGTAAGCCA 253

QY 242 CCTCAGGAGAGAACGTGTTCTTCTTGTTGACCAACTTCTCTGTGACGCCAGCCCAAGTT 301
DB 254 CCTCA----- 258

QY 302 CAGGCAGATCCAGAGACCCCTCGTCCACTGGCTAACTGCTGGTGGACGAGGAC 361
DB 259 ----- 258

QY 362 TGCCCCGAAGGGAGGAGGCACACAGCCAGCTGTAAAGACAGGCCAGTGTGTGGTG 421
DB 259 -----GGTGTAAAGACAGGCCAGTGTGTGGTG 285

QY 422 TTCAATGGGACCCAGAGACCTGTGAGATCTGGAGTTGGTGCCCGAGTGGAGAGTGGCGTT 481
DB 286 TTCAATGGGACCCAGAGACCTGTGAGATCTGGAGTTGGTGCCCGAGTGGAGAGTGGCGTT 345

QY 482 GTGCCCTCAGGCGCCCTCTGCGCCAGCCAGCCAGAACTTCACTGTTTCATCAAAACACA 541
DB 346 GTGCCCTCAGGCGCCCTCTGCGCCAGCCAGCCAGAACTTCACTGTTTCATCAAAACACA 405

QY 542 GTCACCTTCAGCAAGTTCACATTTCTTAAGTTCGAATGCTTGGAGACTGGGACCCCAACC 601
DB 406 GTCACCTTCAGCAAGTTCACATTTCTTAAGTTCGAATGCTTGGAGACTGGGACCCCAACC 465

QY 602 TATTTTAGCACTCCGCTATGACCACTTCACTAGCCCTACTCTCCCGTGTTCGGCAT 661
DB 466 TATTTTAGCACTCCGCTATGACCACTTCACTAGCCCTACTCTCCCGTGTTCGGCAT 525

QY 662 GGGACCTCTGCGGCAAGCTGGAGGACCTTCGAGGACCTGGCGTTGCTGGGTGGCTCT 721
DB 526 GGGACCTCTGCGGCAAGCTGGAGGACCTTCGAGGACCTGGCGTTGCTGGGTGGCTCT 585

QY 722 GTAGGCATCAGATTCACTGGATGTGACCTGGACACCGGGGACTCTGCTCTGCGCT 781
DB 586 GTAGGCATCAGATTCACTGGATGTGACCTGGACACCGGGGACTCTGCTCTGCGCT 645

```

```

QY 782 CACTACTCTCTCCAGCTGCAGGAGAGCTACAACTTCAGGACAGCCACTCACTGGTGG 841
DB 646 CACTACTCTCTCCAGCTGCAGGAGAGCTACAACTTCAGGACAGCCACTCACTGGTGG 705

QY 842 GAGCAACCGGGTGTGGAGGCCCGCACC 868
DB 706 GAGCAACCGGGTGTGGAGGCCCGCACC 732

CFI31935 718 bp mRNA linear EST 05-AUG-2003
UI-HF-FQ0-aws-m-11-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
IMAGE:30553930 5', mRNA sequence.
CFI31935
CFI31935.1 GI:33214691
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 718)
Ronald, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Mary Hendrix
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.
Location/Qualifiers
1. 718
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30553930"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_215"
/notes="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is GATAAGGCCA. Tissue was provided by Mary Hendrix."

FEATURES
source
Query Match 20.5%; Score 553; DB 6; Length 718;
Best Local Similarity 99.5%; Pred. No. 1.5e-121;
Matches 564; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 97 GTGGGCTCTCTCTCGCAAAAGGCTACAGAGCGGAGCTGGAAACCCAGTTTCCAT 156
DB 153 GTGGGCTCTCTCTCGCAAAAGGCTACAGAGCGGAGCTGGAAACCCAGTTTCCAT 212

ORIGIN
Query Match 20.5%; Score 553; DB 6; Length 718;
Best Local Similarity 99.5%; Pred. No. 1.5e-121;
Matches 564; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 97 GTGGGCTCTCTCTCGCAAAAGGCTACAGAGCGGAGCTGGAAACCCAGTTTCCAT 156
DB 153 GTGGGCTCTCTCTCGCAAAAGGCTACAGAGCGGAGCTGGAAACCCAGTTTCCAT 212

```

QY 157 CATCAACAACTCAAGGGGTTTCGTCACCTCAGATCAAGGAGCTTGAAACCGGCTGTG 216  
Db 213 CATCAACAACTCAAGGGGTTTCGTCACCTCAGATCAAGGAGCTTGAAACCGGCTGTG 272  
QY 217 GATGTGGCGGACTTCGTGAAGCCACTCAGGAGAGAACGTGTTCTTCTTGGTGACCAA 276  
Db 273 GATGTGGCGGACTTCGTGAAGCCACTCAGGAGAGAACGTGTTCTTCTTGGTGACCAA 332  
QY 277 CTTCTCTGTGAGCCAGCCAGGCTTCAAGGAGAGATGCCAGAGCACCCTGCTGCCACT 336  
Db 333 CTTCTCTGTGAGCCAGCCAGGCTTCAAGGAGAGATGCCAGAGCACCCTGCTGCCACT 392  
QY 337 GCGTAACCTCTGGTTCGACGAGACTGCCCGAAGGGAGGAGCACAACAGCCAGCG 396  
Db 393 GCGTAACCTCTGGTTCGACGAGACTGCCCGAAGGGAGGAGCACAACAGCCAGCG 452  
QY 397 TGTAAACAGCCAGCTGTGTGTTCAATGGGACCCACAGGACCTGTGATCTGGAG 456  
Db 453 TGTAAACAGCCAGCTGTGTGTTCAATGGGACCCACAGGACCTGTGATCTGGAG 512  
QY 457 TTGGTGGCCAGTGGAGAGTGGGTTGTGCCCTCGAGGCCCCCTGCTGGCCAGGCCAGAA 516  
Db 513 TTGGTGGCCAGTGGAGAGTGGGTTGTGCCCTCGAGGCCCCCTGCTGGCCAGGCCAGAA 572  
QY 517 CTTCACTGTTCATCAAAACACACAGTCACTTCAGAGTTCAACTTCTTAAGTCAA 576  
Db 573 CTTCACTGTTCATCAAAACACACAGTCACTTCAGAGTTCAACTTCTTAAGTCAA 632  
QY 577 TGCCTTGAGAGACTGGGACCCCACTATTTTAAGCACTGCCCTATGAACACCAATTCCAG 636  
Db 633 TGCCTTGAGAGACTGGGACCCCACTATTTTAAGCACTGCCCTATGAACACCAATTCCAG 691  
QY 637 CCCCTACTGTCCCGTGTCCGCAATGG 663  
Db 692 CCCCTACTGTCCCGTGTCCGCAATGG 718

RESULT 12  
BM696978  
LOCUS  
DEFINITION  
UI-E-DW0-agl-f-08-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone  
ACCESSION  
BM696978  
VERSION  
BM696978.1 GI:19010236  
KEYWORDS  
EST.  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 552)  
Ronald, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEERF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@iowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Reverse.  
FEATURES  
Location/Qualifiers  
1..552  
/organism="Homo sapiens"

/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-DW0-agl-f-08-0-UI"  
/tissue\_type="lens"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-DW0"  
/notes="Organ: eye; Vector: pVT3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-E-DW0 is a cDNA library containing the following  
tissue(s): lens. The library was constructed according to  
Bonald, Lennon and Soares, Genome Research, 6:791-806,  
1996. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was ligated to an EcoR I adaptor, digested with Not  
I, and cloned directionally into pVT3-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is CGATTACGA. This library  
was created for the program, Gene Discovery in the Visual  
System, supported by National Eye Institute (NEI)."  
ORIGIN  
Query Match 20.5%; Score 552; DB 4; Length 552;  
Best Local Similarity 100.0%; Pred. No. 2.5e-121;  
Matches 552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1904 AGTTATAGAGACGGCTGCTTCCAGTGTAGCCCTTTTGGCATGGAGCTCTGGGAGAG 1963  
Db 1 AGTTATAGAGACGGCTGCTTCCAGTGTAGCCCTTTTGGCATGGAGCTCTGGGAGAG 60  
QY 1964 AGAGCAGAGGCGGCGGCTAAAGTTGGTGATCATTTGGGTCTTTCAGGACCTTCTATATC 2023  
Db 61 AGAGCAGAGGCGGCGGCTAAAGTTGGTGATCATTTGGGTCTTTCAGGACCTTCTATATC 120  
QY 2024 CTTCTCGGTAAACCCCGCCCAACCCCTTGAATCTTCTCCAGGCTTCTCGAGAGC 2083  
Db 121 CTTCTCGGTAAACCCCGCCCAACCCCTTGAATCTTCTCCAGGCTTCTCGAGAGC 180  
QY 2084 CTTGGGGGTGGGAGGCTGTGGGAGGCTGTACATCTGAAATTCACCTTCAGTCCAAGTCATA 2143  
Db 181 CTTGGGGGTGGGAGGCTGTGGGAGGCTGTACATCTGAAATTCACCTTCAGTCCAAGTCATA 240  
QY 2144 CTTAGGAGCTGTCTGGGAGCTGTCTGAGGAGGCGCTGCTGCTGATCCAGGCTGGAT 2203  
Db 241 CTTAGGAGCTGTCTGGGAGCTGTCTGAGGAGGCGCTGCTGCTGATCCAGGCTGGAT 300  
QY 2204 GGAGTGGCTGGAAGGAATGTTCCAAACACACCCAGAGATCTCCTCAGGCTGGCCAG 2263  
Db 301 GGAGTGGCTGGAAGGAATGTTCCAAACACACCCAGAGATCTCCTCAGGCTGGCCAG 360  
QY 2264 GTTTTCAGCTGGAATTTCTCTTCTTGGTCCAGGGCGGGGAGGAAATTTCTAAGTGTCCA 2323  
Db 361 GTTTTCAGCTGGAATTTCTCTTGGTCCAGGGCGGGGAGGAAATTTCTAAGTGTCCA 420  
QY 2324 CCCCAGGAGGCGAAGGGCTGCTTCCACTGTGGTACCTGTGTGATCAGGCGAGCTGTC 2383  
Db 421 CCCCAGGAGGCGAAGGGCTGCTTCCACTGTGGTACCTGTGTGATCAGGCGAGCTGTC 480  
QY 2384 GAGGCGCAGGGGTGGGGCTGAGACTGGGTGACATCTAGAAATCACCTGCCACCTGGAGCC 2443  
Db 481 GAGGCGCAGGGGTGGGGCTGAGACTGGGTGACATCTAGAAATCACCTGCCACCTGGAGCC 540  
QY 2444 TCAGTAAATGC 2455  
Db 541 TCAGTAAATGC 552  
RESULT 13  
BI757011  
LOCUS  
DEFINITION  
BI757011 1038 bp mRNA linear EST 25-SEP-2001  
603028734F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5198937 5',



1101192-1101959, 1217928-1220615); NCI CGAP\_Co10 pool 1  
 LLAM 2644-2653, 2871-2872 (IMAGE clones)  
 1057416-1061255, 1144584-1145351). (50% of the driver  
 population), plus a pool of 3,840 arrayed clones from  
 NCI CGAP Sub1 (IMAGE clones 2708616-2710535) and  
 NCI CGAP Sub2 (IMAGE clones 2710536-2712455) (20% of  
 the driver population), plus a pool of 11,136 clones from  
 NCI CGAP Sub3 (IMAGE clones 2712456-2723591) (30% of  
 the driver population). Subtraction was performed as  
 previously described [Bonaldo, Lennon & Soares (1996):  
 Normalization and Subtraction: Two Approaches to  
 Facilitate Gene Discovery. Genome Research 6, 791-806.  
 TAG\_TISSUE=brain  
 TAG\_LIB=NCI CGAP\_Brn50  
 TAG\_SEQ=TTTCG"

ORIGIN

Query Match 17.3%; Score 466.4; DB 2; Length 477;  
 Best Local Similarity 98.7%; Pred. No. 8.7e-101;  
 Matches 470; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2190 ATCCAGGCTGGATGAGTGGCTGGAAGAAATGGTTCCAAACCAACACCGAGATCTCC 2249  
 DB 476 ATCCAGGCTGGATGAGTGGCTGGAAGAAATGGTTCCAAACCAACACCGAGATCTCC 417  
 QY 2250 CTCAGGCTGGCCAGGTTTTCAGCTGGAAATTCCTCTTGTGTCAGGCGGGGAGGGA 2309  
 DB 416 CTCAGGCTGGCTCAGGTTTTCAGCTGGAAATTCCTCTTGTGTCAGGCGGGGAGGGA 357  
 QY 2310 ATTCTAAGTGTCCACCCAGGAGCAAGGGGCTCTTCCACTGTGGTACCTGGTGTAT 2369  
 DB 356 ATTCTAAGTGTCCACCCAGGAGCAAGGGGCTCTTCCACTGTGGTACCTGGTGTAT 297  
 QY 2370 CAGGCGCAAGCTGTGGAGGCCAGGGGTGGGCTGAGACTGGGCTGACATCTAGAAATCAC 2429  
 DB 296 CAGGCGCAAGCTGTGGAGGCCAGGGGTGGGCTGAGACTGGGCTGACATCTAGAAATCAC 237  
 QY 2430 TGCCACCTGGAGCTCAGTAAATCCCTGGGTCCTCTGCTCTCTAATCTCCAGAGCC 2489  
 DB 236 TGCCACCTGGAGCTCAGTAAATCCCTGGGTCCTCTGCTCTCTAATCTCCAGAGCC 177  
 QY 2490 ATGTCATGGGAGGCTGGGCTCTGAAGGGCGAAGGTGGGAGAGGAGGCGCCCTCAGGCT 2549  
 DB 176 ATGTCATGGGAGGCTGGGCTCTGAAGGGCGAAGGTGGGAGAGGAGGCGCCCTCAGGCT 117  
 QY 2550 GGGTATCCAGAGGAGGCACTGACCTGATTCCTTTGGGCCCCAGAGGAGCTGATGT 2609  
 DB 116 GGGTATCCAGAGGAGGCACTGACCTGATTCCTTTGGGCCCCAGAGGAGCTGATGT 57  
 QY 2610 CATGCTCGACAAAGTCAAGGAGTAAGCCAGCAAGCCCAAAACCAAAAAA 2665  
 DB 56 CATGCTCGACAAAGTCAAGGAGTAAGCCAGCAAGCCCAAAACCAAAAAA 1

RESULT 15  
 BM689096  
 LOCUS  
 DEFINITION UI-E-Q1-act-a-11-0-UI.r1 UI-E-Q1 Homo sapiens cDNA clone  
 UI-E-Q1-act-a-11-0-UI 5', mRNA sequence.  
 ACCESSION BM689096  
 VERSION BM689096.1 GI:19002354  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 460)  
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (3), 791-806 (1996)  
 MEDLINE 9704477  
 PUBMED 889548

## COMMENT

Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clones Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.

## FEATURES

Location/Qualifiers  
 1..460  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-Q1-act-a-11-0-UI"  
 /tissue\_type="optic nerve"  
 /dev\_stage="adult"  
 /lab\_host="DHI08 (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-Q1"  
 /notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-Q1 is a normalized cDNA library containing the  
 following tissue(s): optic nerve. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT7T3-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dfr)18 tail. The sequence tag for this library is  
 CCAATTAAGT. This library was created for the program, Gene  
 Discovery in the Visual System, supported by National Eye  
 Institute (NEI)."

## ORIGIN

Query Match 16.7%; Score 450.4; DB 4; Length 460;  
 Best Local Similarity 99.8%; Pred. No. 6e-97;  
 Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2070 GGCTTCTCAGAGCCCTGGGGTGGGAGGCTGTGGAGGCTGTACATCTGAAATTCACCT 2129  
 DB 9 GGCTTCTCAGAGCCCTGGGGTGGGAGGCTGTGGAGGCTGTACATCTGAAATTCACCT 68  
 QY 2130 CAGTCAAAGTCATACCTAGAAAGCTGTCTGGGAGCTGCTCGAGGAGGCGCTGGCTCTG 2189  
 DB 69 CAGTCAAAGTCATACCTAGAAAGCTGTCTGGGAGCTGCTCGAGGAGGCGCTGGCTCTG 128  
 QY 2190 ATCCAGGCTGGATGAGTGGCTGGAGAAATGGTTCCAAACCAACACCGAGATCTCC 2249  
 DB 129 ATCCAGGCTGGATGAGTGGCTGGAGAAATGGTTCCAAACCAACACCGAGATCTCC 188  
 QY 2250 CTCAGGCTGGCCAGGTTTTCAGCTGGAAATTCCTCTTGGTCCAGGCGGGGAGGGA 2309  
 DB 189 CTCAGGCTGGCCAGGTTTTCAGCTGGAAATTCCTCTTGGTCCAGGCGGGGAGGGA 248  
 QY 2310 ATTCTAAGTGTCCACCCAGGAGGAGGAGGCTGCTTTCCACTGTGGGTACTGGTGTAT 2369  
 DB 249 ATTCTAAGTGTCCACCCAGGAGGAGGAGGCTGCTTTCCACTGTGGGTACTGGTGTAT 308  
 QY 2370 CAGGGCAAGCTGTGGAGGCCAGGGGTGGGCTGAGACTGGGCTGACATCTAGAAATCAC 2429  
 DB 309 CAGGGCAAGCTGTGGAGGCCAGGGGTGGGCTGAGACTGGGCTGACATCTAGAAATCAC 368  
 QY 2430 TGCCACCTGGAGCTCAGTAAATCCCTGGGTCCTGCTCTCTAATCTCCAGAGCC 2489  
 DB 369 TGCCACCTGGAGCTCAGTAAATCCCTGGGTCCTGCTCTCTAATCTCCAGAGCC 428

Sun Nov 28 09:38:14 2004

Qy 2490 ATGTCCATGGGGAGGTGGGCTCTGAAGGGCGA 2521  
| | | | | | | | | | | | | | | | | | | | | |  
Db 429 ATGTCCATGGGGAGGTGGGCTCTGAAGGGTGA 460

Search completed: November 21, 2004, 17:35:05  
Job time : 8637.1 secs



Result	No.	Score	Query	Match	Length	DB	ID	Description
1		2593	100.0	2693	10	US-09-820-095-1		Sequence 1, Appl
2		1592.6	59.1	16449	10	US-09-820-095-3		Sequence 3, Appl
C	3	567.4	21.1	569	9	US-09-864-761-9190		Sequence 9190, Ap
C	4	377	14.0	577	9	US-09-864-761-9695		Sequence 9695, Ap
5		243.6	9.0	1978	15	US-10-172-118-786		Sequence 786, App
6		243.6	9.0	1978	16	US-10-342-887-786		Sequence 786, App
7		239.2	8.9	1389	9	US-09-833-082-1		Sequence 1, Appl
8		239.2	8.9	1750	16	US-10-305-720-831		Sequence 831, App
9		237.6	8.8	1269	18	US-10-128-558-20		Sequence 20, Appl
10		237.2	8.8	1167	16	US-10-386-414-18		Sequence 18, Appl
11		237.2	8.8	2048	16	US-10-386-414-16		Sequence 16, Appl
12		237.2	8.8	2048	16	US-10-240-423-1468		Sequence 1468, Ap

181 CGTCACTCAGATCAAGAGCTTGAAACCGCGCTGTGGATGTGGCGACTTCTGTGAAGCC 240 QY  
181 CGTCACTCAGATCAAGAGCTTGAAACCGCGCTGTGGATGTGGCGACTTCTGTGAAGCC 240 Db  
241 ACCTCAGGAGAGAACCTGTTCTTCTGTGTGAGCAACTTCTGTGACGCCAGCCAAAGT 300 QY  
241 ACCTCAGGAGAGAACCTGTTCTTCTGTGTGAGCAACTTCTGTGACGCCAGCCAAAGT 300 Db  
301 TCAGGAGAGATGCCAGAGACCGCTGCCCTCCCACTGGCTAACTGTGGGTGCAGAGGA 360 QY  
301 TCAGGAGAGATGCCAGAGACCGCTGCCCTCCCACTGGCTAACTGTGGGTGCAGAGGA 360 Db  
361 CTGCCCCAAGGGAGGAGGAGCACAACAGCAGCGGTGTAACACAGCCAGCTGTGTGT 420 QY  
361 CTGCCCCAAGGGAGGAGGAGCACAACAGCAGCGGTGTAACACAGCCAGCTGTGTGT 420 Db  
421 GTTCAATGGGACCCACAGGACCTGTGAGATCTGGAGTCTGGAGTGTGGCCAGTGGCGT 480 QY  
421 GTTCAATGGGACCCACAGGACCTGTGAGATCTGGAGTCTGGAGTGTGGCCAGTGGCGT 480 Db  
481 TGTCCCTCGAGGCCCCGTGTGGCCAGGCCAGAACTTCAACTGTTCATCAAAAACAC 540 QY  
481 TGTCCCTCGAGGCCCCGTGTGGCCAGGCCAGAACTTCAACTGTTCATCAAAAACAC 540 Db  
541 AGTCACCTTCAGCAAGTCAACTTCTTAAGTCCATGCTTGGAGACCTGGGACCCAC 600 QY  
541 AGTCACCTTCAGCAAGTCAACTTCTTAAGTCCATGCTTGGAGACCTGGGACCCAC 600 Db  
601 CTATTTAAGCACTGCCGTATGAACCAAACTTACGCCCTTACGTCCCGTGTCCGCT 660 QY  
601 CTATTTAAGCACTGCCGTATGAACCAAACTTACGCCCTTACGTCCCGTGTCCGCT 660 Db  
661 TGGGACCTGTGGCCAAAGCTGAGGAGCCTTCGAGGACCTGGCGTGTGGTGGCTC 720 QY  
661 TGGGACCTGTGGCCAAAGCTGAGGAGCCTTCGAGGACCTGGCGTGTGGTGGCTC 720 Db  
721 TGTAGGACATCAGAGTTCACTGGGATGTGACCTGGACACCGGGACTCTGGCTGTGGCC 780 QY  
721 TGTAGGACATCAGAGTTCACTGGGATGTGACCTGGACACCGGGACTCTGGCTGTGGCC 780 Db  
781 TCACCTACTCTCCAGCTGAGGAGAGAGTACAACTTCAGGACGCCACTCACTGGTG 840 QY  
781 TCACCTACTCTCCAGCTGAGGAGAGAGTACAACTTCAGGACGCCACTCACTGGTG 840 Db  
841 GGAGCAACCGGGTGTGGAGGCCCGACCTGTCAAGCTCTATGGAATCCGCTTCGACAT 900 QY  
841 GGAGCAACCGGGTGTGGAGGCCCGACCTGTCAAGCTCTATGGAATCCGCTTCGACAT 900 Db  
901 CCTCGTCAACCGGGAGGAGAGTTCGGGCTCATCCACGCCCGTCACTGGGCGAC 960 QY  
901 CCTCGTCAACCGGGAGGAGAGTTCGGGCTCATCCACGCCCGTCACTGGGCGAC 960 Db  
961 CGGGGAGCTTGGCTGGCGGTGTCACCTTTCTGTGACCTCTACTGTGTGTGTGGA 1020 QY  
961 CGGGGAGCTTGGCTGGCGGTGTCACCTTTCTGTGACCTCTACTGTGTGTGTGGA 1020 Db  
1021 TAGAGAGCCCATTTCTACTGGAGGACAAAGTATGAGGAGGCCAAGGCCCGGAAAGCAAC 1080 QY  
1021 TAGAGAGCCCATTTCTACTGGAGGACAAAGTATGAGGAGGCCAAGGCCCGGAAAGCAAC 1080 Db  
1081 CGCAACTCTGTGTGGAGGAGCTGGCGCTTGCATCCAGCCCGACTGGCCAGTGCCT 1140 QY  
1081 CGCAACTCTGTGTGGAGGAGCTGGCGCTTGCATCCAGCCCGACTGGCCAGTGCCT 1140 Db  
1141 CAGACGGAGCTCAGACCTGACCCACCGCCACTGCTGCTGGAGTCAGACACAGACACC 1200 QY  
1141 CAGACGGAGCTCAGACCTGACCCACCGCCACTGCTGCTGGAGTCAGACACAGACACC 1200 Db  
1201 AGGATGGCCCTGTCAAGTTCGACCCACTTGCACACCCATTCGGGAGCTGTAGCC 1260 QY  
1201 AGGATGGCCCTGTCAAGTTCGACCCACTTGCACACCCATTCGGGAGCTGTAGCC 1260 Db

1261 GTTCCCTCTGCTGGTTGAGAGTTGGGGCTGGGAAGGGGGGCCCTGCTGGGATCTCAA 1320 QY  
1261 GTTCCCTCTGCTGGTTGAGAGTTGGGGCTGGGAAGGGGGGCCCTGCTGGGATCTCAA 1320 Db  
1321 GGATGAGCCCCCAGCATGAGGATTTGGGGTAGAATTTCCACCTTGAACCCAGCAGACA 1380 QY  
1321 GGATGAGCCCCCAGCATGAGGATTTGGGGTAGAATTTCCACCTTGAACCCAGCAGACA 1380 Db  
1381 GTCCCTCCCTGACTCCACCTTGTAGGGTGTGCTCAGGGAGCCATGAAATCGGCT 1440 QY  
1381 GTCCCTCCCTGACTCCACCTTGTAGGGTGTGCTCAGGGAGCCATGAAATCGGCT 1440 Db  
1441 GTGTTTTGAGACGGCGACAGAACTTGACCCCGTGAGAGCTGGGAGAGCCAGCAGCACCT 1500 QY  
1441 GTGTTTTGAGACGGCGACAGAACTTGACCCCGTGAGAGCTGGGAGAGCCAGCAGCACCT 1500 Db  
1501 GTATTGCAAGGCTCCGACTGTCATGTGGCAGGGGTCTGTGCTGGCTCTGGGCTGGAGTC 1560 QY  
1501 GTATTGCAAGGCTCCGACTGTCATGTGGCAGGGGTCTGTGCTGGCTCTGGGCTGGAGTC 1560 Db  
1561 TCTCTCCAGTGTCTGTGCCAGTGTTCCTAGCAGAGGTATGCTTACCAGCTGTCAAGCA 1620 QY  
1561 TCTCTCCAGTGTCTGTGCCAGTGTTCCTAGCAGAGGTATGCTTACCAGCTGTCAAGCA 1620 Db  
1621 CAGACCTCTCTGCTGCTGGGTCTGGCCCTCTCCGCCATCTGCACCCCATCATAGGT 1680 QY  
1621 CAGACCTCTCTGCTGCTGGGTCTGGCCCTCTCCGCCATCTGCACCCCATCATAGGT 1680 Db  
1681 AGAGACCCACCTCCCATCGTCTCATCGGGCTGTGACAGCTGGAGCCAAAAGGCAA 1740 QY  
1681 AGAGACCCACCTCCCATCGTCTCATCGGGCTGTGACAGCTGGAGCCAAAAGGCAA 1740 Db  
1741 GGCAGAAAGAGAGTGTATGGGGAGGGGATTTGTTTCACTTCTCTGCTGTGTATGCC 1800 QY  
1741 GGCAGAAAGAGAGTGTATGGGGAGGGGATTTGTTTCACTTCTCTGCTGTGTATGCC 1800 Db  
1801 CCAGAGAGTCTTAATCTAGGGAATGGGTGGAGTAGGAGATATCCATCCCTCTATCC 1860 QY  
1801 CCAGAGAGTCTTAATCTAGGGAATGGGTGGAGTAGGAGATATCCATCCCTCTATCC 1860 Db  
1861 CCAGGCAAGGGCGGAGCATGTGCTTTGGGCCACACCTGCTTAGTTATGAGGACCGGC 1920 QY  
1861 CCAGGCAAGGGCGGAGCATGTGCTTTGGGCCACACCTGCTTAGTTATGAGGACCGGC 1920 Db  
1921 TGCTTTCAGTGGTAGCCCTTTTGCCATGAGAGTCTGGGAGAGAGAGGCGGCGAG 1980 QY  
1921 TGCTTTCAGTGGTAGCCCTTTTGCCATGAGAGTCTGGGAGAGAGAGGCGGCGAG 1980 Db  
1981 GGCTAAGTTGGTATCATTTGGGTCTTCAGGACCTTCTATATCCCTCTCGGTAAACCCC 2040 QY  
1981 GGCTAAGTTGGTATCATTTGGGTCTTCAGGACCTTCTATATCCCTCTCGGTAAACCCC 2040 Db  
2041 CAGCCCAACCCCTTGGAACTTTCTCCAGGCTTCTGAGAGCCCTGGGGTGGAGGCT 2100 QY  
2041 CAGCCCAACCCCTTGGAACTTTCTCCAGGCTTCTGAGAGCCCTGGGGTGGAGGCT 2100 Db  
2101 GTGGAGAGCTGTACATCTGAAATTCACCTTCACTTCACTAGTCCAGTATACCTAGGAGCTGTGG 2160 QY  
2101 GTGGAGAGCTGTACATCTGAAATTCACCTTCACTTCACTAGTCCAGTATACCTAGGAGCTGTGG 2160 Db  
2161 GCAGTGTCTCAGGAGGAGCCCTGCTCTGATCCAGGCTGGATGGATGGCTGGAGGAA 2220 QY  
2161 GCAGTGTCTCAGGAGGAGCCCTGCTCTGATCCAGGCTGGATGGATGGCTGGAGGAA 2220 Db  
2221 TGGTTCAAAACAAACCCAGAGATCTCCCTCAGGCTGGCCAGGTTTTCAGCTGGAAAT 2280 QY  
2221 TGGTTCAAAACAAACCCAGAGATCTCCCTCAGGCTGGCCAGGTTTTCAGCTGGAAAT 2280 Db  
2281 CTCTCTTTGGTCCCAAGGGGGGAGGAAATTTAAAGTGTCCACCCAGGAGGACAAAGG 2340 QY  
2281 CTCTCTTTGGTCCCAAGGGGGGAGGAAATTTAAAGTGTCCACCCAGGAGGACAAAGG 2340 Db  
2341 GCTGTCTTCCACTGTGGGTACTGTGGTATCAGGGCAAGCTGTGGAGGCGCAGGGGTGGGG 2400 QY

```
Db 2341 GCTGCTTCCACTGTGGTACCTGGTGATCAGGCAAGCTGTGGAGGCCAGGGGTGGG 2400
QY 2401 CTGAGACTGGGCTGACATCTAGATACCTGCGACCTGGAGCTCAGTAAATCCCTGGG 2460
Db 2401 CTGAGACTGGGCTGACATCTAGATACCTGCGACCTGGAGCTCAGTAAATCCCTGGG 2460
QY 2461 GTCCTGTGCTCTCAATCTCCAGAGCCATGTCCATGGGAGGTGGGCTCTGAAGGGCG 2520
Db 2461 GTCCTGTGCTCTCAATCTCCAGAGCCATGTCCATGGGAGGTGGGCTCTGAAGGGCG 2520
QY 2521 AAGTGGAGAGCAGGGCCCTGAGGCTGGGATCCAGAGGGGCGACGTGCACTGAT 2580
Db 2521 AAGTGGAGAGCAGGGCCCTGAGGCTGGGATCCAGAGGGGCGACGTGCACTGAT 2580
QY 2581 TCTCCTTGGGCCCAGAGGAGAGCTGATGCTATGCTGGACAAAGTTCACGAGTAAAGCCA 2640
Db 2581 TCTCCTTGGGCCCAGAGGAGAGCTGATGCTATGCTGGACAAAGTTCACGAGTAAAGCCA 2640
QY 2641 GCAAAGCCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2693
Db 2641 GCAAAGCCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2693

RESULT 2
US-09-820-095-3
; Sequence 3, Application US/09820095
; Publication No. US20030233668A1
; GENERAL INFORMATION:
; APPLICANT: WEI Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001202
; CURRENT APPLICATION NUMBER: US/09/820,095
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16449
; TYPE: DNA
; ORGANISM: Human
US-09-820-095-3

Query Match 59.1%; Score 1592.6; DB 10; Length 16449;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1595; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1053 ATGAGGAGCCAAAGGCCCGGAAAGCAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTG 1112
Db 13246 ATCTGCAGCCCAAGGCCCGGAAAGCAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTG 13305
QY 1113 CATCCCAAGCCGACTGGCCGAGTGCTCTCAGCGGAGCTCAGACCTGACCCACCGGCCA 1172
Db 13306 CATCCCAAGCCGACTGGCCGAGTGCTCTCAGCGGAGCTCAGACCTGACCCACCGGCCA 13365
QY 1173 CTGCTGCTGGGAGTCAGACAGACACAGATGSCCTGTCGAAGTTCTGACACCCACT 1232
Db 13366 CTGCTGCTGGGAGTCAGACAGACACAGATGSCCTGTCGAAGTTCTGACACCCACT 13425
QY 1233 TGCCAAACCAATTCGGGAGCTGTAGCCGTTCCCTGTGTTGAGAGTTGGGGGCTGGGA 1292
Db 13426 TGCCAAACCAATTCGGGAGCTGTAGCCGTTCCCTGTGTTGAGAGTTGGGGGCTGGGA 13485
QY 1293 AGGGCGGGGCTGCTGCTGGGATCTCAAGGATGAGGCCCGCCAGCATGGAGGATTGGGGTA 1352
Db 13486 AGGGCGGGGCTGCTGCTGGGATCTCAAGGATGAGGCCCGCCAGCATGGAGGATTGGGGTA 13545
QY 1353 GAATTCACCCCTTGAACCCCGACAGACAGTCCCTCCCTGACTCCCACTTGGTAGGGTG 1412
Db 13546 GAATTCACCCCTTGAACCCCGACAGACAGTCCCTCCCTGACTCCCACTTGGTAGGGTG 13605
QY 1413 CTGCTCAGGAGGCCATAGAGTGGGCTGTGTTTGGAGCGGCGACAGAACTGACCCGT 1472
```

```
Db 13606 CTGCTCAGGAGCCATAGAGTCGGCTGTGTTTGGAGCGCGACAGAACTGACCCGT 13665
QY 1473 GGAGACTGGAGAGCCCGACAGCAGCAGCTGTATTGGAGGCTCCGACTGCTGATGGCAGG 1532
Db 13666 GGAGACTGGAGAGCCCGACAGCAGCAGCTGTATTGGAGGCTCCGACTGCTGATGGCAGG 13725
QY 1533 GTCCTGTGCTGCTGTGGGCTGGAGTCTCTCTCCAGTGTCTGTGCCAGTGTTCCTA 1592
Db 13726 GTCCTGTGCTGCTGTGGGCTGGAGTCTCTCTCCAGTGTCTGTGCCAGTGTTCCTA 13785
QY 1593 GCAGAGTATGCTTACCAGCTGTGACGACAGACCCCTCTGCTGCTGGGTCTGGCCCTC 1652
Db 13786 GCAGAGTATGCTTACCAGCTGTGACGACAGACCCCTCTGCTGCTGGGTCTGGCCCTC 13845
QY 1653 CTCGCCCATCTGCACCCCATCATAGGTAGACACCCACCCCTCCCATCCGTCCTACATGG 1712
Db 13846 CTCGCCCATCTGCACCCCATCATAGGTAGACACCCACCCCTCCCATCCGTCCTACATGG 13905
QY 1713 GGTGTGCACTGTGGAGCCCAAGGCAAGGCAAGAGAGTGTATGGGGAGGGGGATT 1772
Db 13906 GGTGTGCACTGTGGAGCCCAAGGCAAGGCAAGAGAGTGTATGGGGAGGGGGATT 13965
QY 1773 GTTTCAGCTTCTGCTGCTGTATGCCCGACGAGATCCTAATCTAGGGAATGGGTGG 1832
Db 13966 GTTTCAGCTTCTGCTGCTGTATGCCCGACGAGATCCTAATCTAGGGAATGGGTGG 14025
QY 1833 AGTAGGCAGATATCCACCTCCCTATCCCCAGGCAAGCGGAGCATGTGTGGGCC 1892
Db 14026 AGTAGGCAGATATCCACCTCCCTATCCCCAGGCAAGCGGAGCATGTGTGGGCC 14085
QY 1893 CACACCTGCTAGTTATGAGGACCGCTCTTCCAGTGGTAGCCCTTTCCCATGGAG 1952
Db 14086 CACACCTGCTAGTTATGAGGACCGCTCTTCCAGTGGTAGCCCTTTCCCATGGAG 14145
QY 1953 GTCTGGGAGAGAGAGCAGAGGCGCGCAGGGCTAAAGTTGGTGATCATTTGGGTCTTC 2012
Db 14146 GTCTGGGAGAGAGAGCAGAGGCGCGCAGGGCTAAAGTTGGTGATCATTTGGGTCTTC 14205
QY 2013 CTTCTATATTCCTCTCGGTAAACCCCGACGCCAACCCCTTGGATCTTCTCCACAGC 2072
Db 14206 CTTCTATATTCCTCTCGGTAAACCCCGACGCCAACCCCTTGGATCTTCTCCACAGC 14265
QY 2073 TTCTCAGAGCCCTGGGGTGGAGGCTGTGGAGGCTGTACATCTGAAATTCACCTTCAG 2132
Db 14266 TTCTCAGAGCCCTGGGGTGGAGGCTGTGGAGGCTGTACATCTGAAATTCACCTTCAG 14325
QY 2133 TCCAAGTCTACCTAGGAAAGTGTCTGGGAGCTGCTCAGGGAGGCCCTGGCTCTGATC 2192
Db 14326 TCCAAGTCTACCTAGGAAAGTGTCTGGGAGCTGCTCAGGGAGGCCCTGGCTCTGATC 14385
QY 2193 CAGGCTGGATGGAGTGGCTGGAGGAATGGTTCCAAACAAACACACCCAGATCTCCCTC 2252
Db 14386 CAGGCTGGATGGAGTGGCTGGAGGAATGGTTCCAAACAAACACACCCAGATCTCCCTC 14445
QY 2253 AGGCTGGCAGAGTTTTCAGCTGGAATCTCTCTTGTGTCAGGGCGGGGAGGGAAT 2312
Db 14446 AGGCTGGCAGAGTTTTCAGCTGGAATCTCTCTTGTGTCAGGGCGGGGAGGGAAT 14505
QY 2313 CTAAGTGTCCACCCAGAGGAGGCAAGGGCTGCTTCCACTGTGGGTACTGCTGATCAG 2372
Db 14506 CTAAGTGTCCACCCAGAGGAGGCAAGGGCTGCTTCCACTGTGGGTACTGCTGATCAG 14565
QY 2373 GGCAAGCTGTGGAGGCGCCAGGGGTGGGCTGAGACTGGGTGACATCTAGAAATCACCTGC 2432
Db 14566 GGCAAGCTGTGGAGGCGCCAGGGGTGGGCTGAGACTGGGTGACATCTAGAAATCACCTGC 14625
QY 2433 CACCTGGAGCCTCAGTAAATAGCTGGGTCCCTGTGCTCTCAATCTCCAGAGCCATG 2492
Db 14626 CACCTGGAGCCTCAGTAAATAGCTGGGTCCCTGTGCTCTCAATCTCCAGAGCCATG 14685
QY 2493 TCCATGGGAGGTGGGCTCTGAGGCGGAAGTGGGAGAGCGGGCCCTGAGGCGCTGGG 2552
```

Db 14686 TCCATGGGAGGTGGCTCTGAAGGCGAAGGTGGGAGAGAGAGGCCCCCTGAGGCCTGGG 14745  
QY 2553 TATCCAGAGGGGACGTCACCTGATTCCTTTGGGGCCAGAGGAAGCTGATGTCAT 2612  
Db 14746 TATCCAGAGGGGACGTCACCTGATTCCTTTGGGGCCAGAGGAAGCTGATGTCAT 14805  
QY 2613 GCCTGGCAAAAGTCACGAGTAAAGCCAGCAAAAGCCACC 2651  
Db 14806 GCCTGGCAAAAGTCACGAGTAAAGCCAGCAAAAGCCACC 14844

RESULT 3  
US-09-864-761-9190/c  
; Sequence 9190, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecmca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,697  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ IDS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 9190  
; LENGTH: 569  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC002472.3  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4  
US-09-864-761-9190

Query Match 21.1%; Score 567.4; DB 9; Length 569;  
Best Local Similarity 99.8%; Pred. No. 9.2e-14; Indels 0; Gaps 0;  
Matches 568; Conservative 0; Mismatches 14;

QY 1497 ACCTGTATTGAGGGCTCCGACTGTCATGTGCGAGGGCTCCTGCTGGGCTTGGGCTTGGG 1556  
Db 569 ACCTGTATTGAGGGCTCCGACTGTCATGTGCGAGGGCTCCTGCTGGGCTTGGGCTTGGG 510  
QY 1557 GGTCTCTCTCCAGTGTCTGTGCCAGTGTCTTAGCAGAGGTATGCTTACCAAGCTGTC 1616  
Db 509 GGTCTCTCTCCAGTGTCTGTGCCAGTGTCTTAGCAGAGGTATGCTTACCAAGCTGTC 450  
QY 1617 AGCACAGACCCCTCCTGCTGCTGGTCTGCTGCCCTCCTCCTCCCATCTGCACCCCATCAT 1676  
Db 449 AGCACAGACCCCTCCTGCTGCTGGTCTGCTGCCCTCCTCCTCCCATCTGCACCCCATCAT 390  
QY 1677 AGGTAGAGACCCCACTCCCATCGGTCCTACATGGGGCTGTGAGCTGGAGGCCAAAAG 1736  
Db 389 AGGTAGAGACCCCACTCCCATCGGTCCTACATGGGGCTGTGAGCTGGAGGCCAAAAG 330  
QY 1737 GCAGGCGAAGAGGAGTGTGGGGAGGGGATTTTCAGCTTCTCTGGTCTGTGA 1796  
Db 329 GCAGGCGAAGAGGAGTGTGGGGAGGGGATTTTCAGCTTCTCTGGTCTGTGA 270  
QY 1797 TGCCCCAGGAGTCTTAATCTAGGGAATGGGGTGGAGTAGGAGAGATAATCCACTCCCT 1856  
Db 269 TGCCCCAGGAGTCTTAATCTAGGGAATGGGGTGGAGTAGGAGAGATAATCCACTCCCT 210  
QY 1857 ATCCCCAGGCAAGGGCGGAGCATGTCTTTGGGGCCACACCTCTTAGTTATGAGGAC 1916  
Db 209 ATCCCCAGGCAAGGGCGGAGCATGTCTTTGGGGCCACACCTCTTAGTTATGAGGAC 150  
QY 1917 CGGCTGCTTTCCAGTGTAGCCCTTTTGCCATGAGGTCTGGGAGAGAGAGAGGGCG 1976  
Db 149 CGGCTGCTTTCCAGTGTAGCCCTTTTGCCATGAGGTCTGGGAGAGAGAGAGGGCG 90  
QY 1977 GCAGGCTTAAGTGTGATCATTTGGGTCTTCAGGACCTTCTATATATCCCTCTCGGTAA 2036  
Db 89 GCAGGCTTAAGTGTGATCATTTGGGTCTTCAGGACCTTCTATATATCCCTCTCGGTAA 30  
QY 2037 CCCCCAGGCCAACCCCTTGAATCTTTCC 2065  
Db 29 CCCCCAGGCCAACCCCTTGAATCTTTCC 1

RESULT 4  
US-09-864-761-9695/c  
; Sequence 9695, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecmca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04

```

; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/508,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9695
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002472.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 15
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 47
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 17
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 59
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 25
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 22
; US-09-864-761-9695

Query Match          14.0%; Score 377; DB 9; Length 577;
Best Local Similarity 100.0%; Pred. No. 1.4e-94;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2275 GGAATTCCTCTTGGTCCCGGGGGGGGCGGAGGAAATCTAAGTGTCCACCCCGGGAGG 2334
Db 577 GGAATTCCTCTTGGTCCCGGGGGGGGCGGAGGAAATCTAAGTGTCCACCCCGGGAGG 518

Qy 2335 CAAGGGGTGCTTCCACTGTGGTACCTGTGATCAGGGCAAGCTGTGGAGGCCAGGG 2394
Db 517 CAAGGGGTGCTTTCCTCACTGTGGTACCTGTGATCAGGGCAAGCTGTGGAGGCCAGGG 458

Qy 2395 GTGGGGCTGAGACTGGGTGACATCTAGAATCACCTGCCACCTGAGGCTCAGTAAATG 2454
Db 457 GTGGGGCTGAGACTGGGTGACATCTAGAATCACCTGCCACCTGAGGCTCAGTAAATG 398

Qy 2455 CTTGGGTCCCTGCTGCTCTCAATCTCCAGAGCCATGCTCATGGGAGGTGGGCTCTGA 2514
Db 397 CTTGGGTCCCTGCTGCTCTCAATCTCCAGAGCCATGCTCATGGGAGGTGGGCTCTGA 338

Qy 2515 AGGGCGAAGGTGGGAGAGCAGGGCCCTTGAAGCCCTGGGTATCCAAAGAGGGGCCAGTGCA 2574
Db 337 AGGGCGAAGGTGGGAGAGCAGGGCCCTTGAAGCCCTGGGTATCCAAAGAGGGGCCAGTGCA 278

Qy 2575 CTTGATTTCTCTTGGGGCCAGAGAACTGATGTATGTGTGACAAAGTCAACGAGTA 2634
Db 277 CTTGATTTCTCTTGGGGCCAGAGAACTGATGTATGTGTGACAAAGTCAACGAGTA 218
```

```

Qy 2635 AGCCAGCAAAAGCCACC 2651
Db 217 AGCCAGCAAAAGCCACC 201

RESULT 5
US-10-172-118-786
; Sequence 786, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 786
; LENGTH: 1978
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 002561
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-786
```

```

Query Match          9.0%; Score 243.6; DB 15; Length 1978;
Best Local Similarity 58.3%; Pred. No. 3.9e-57;
Matches 501; Conservative 0; Mismatches 329; Indels 30; Gaps 3;

Qy 98 TGGGTCTCTCTCGCCAAAAGGCTACAGGAGCGGGAGCTGGAACCCAGATTTTCCATC 157
Db 175 TGGGTCTCTCTGATAAAGAGGGTTTACCAAGAGCTCGACACCTCTGAGAGTGTGTC 234

Qy 158 ATCAACAAATCAAAAGGGTTTCCGTCATCTAGATCAAGAGCTTGGAAACCGGCTGTG 217
Db 235 ATCAACAAATCAAGGGCGTGGCTTTCACCAACACCTCGGATCTTGGGAGGGATCTGG 294

Qy 218 GATGTGCGGACTTCGTGAAGCCACCTCAGGGAGAGAACTGTCTTCTTGGTACCAAC 277
Db 295 GATGTGCGGACTTCGTGAAGCCACCTCAGGGAGAGAACTGTCTTCTTGGTACCAAC 354

Qy 278 TTCCTTTGTGACGCCAGCCCAAGTTCAAGGCGAGATGCCAGAGCACCCGCTCCCACTG 337
Db 355 CTGATTGTGACCCCAACCCAGCGGAGAACTGTGTGCTGAGAAATGAAGGCATTCCTGAT 414

Qy 338 GCTAACTGTGGTCCAGGAGACTGCCCGAAGGGAGGGAGGCACACACAGCCAGGT 397
Db 415 GCGCGTGTCTCAAGGAGACGCGACTGCCACGCTGGGGAAGCGTTACAGCTGGAAACGGA 474

Qy 398 GTAAAAACAGGCCAGTGTGTGGTG---TTCAATGGGACCCACAGGACCTGTGAGATCTGG 454
Db 475 GTCAAGACGCGCGCTGCTTGGGAGAGGAACTTGGCCAGGGGACCTGTGAGATCTTT 534

Qy 455 AGTTGTGCGCCAGTGAGAGTGGCGTTGTGCCCTGAGGCCCTGTGTCGCCAGGCCAG 514
Db 535 GCCTGTGCGCGTGGAGACAAGCTCCAGGCCGGAGAGCCATTCCTGAAGGAGGCCGAA 594

Qy 515 AACTTCACACTGTTTCATCAAAACACAGTCACCTTCAGCAAGTTCACTTCTTAAGTCC 574
Db 595 GATTCACCAATTTTCATAAAGAACCACTCCGTTTCCCAAAATTAACCTTCTCCAAAC 654

Qy 575 AATGCTCTGGAGACCTTGGGACCCCACTATTTTAAGCACTGGCGCTATGAACCAATTC 634
Db 575 AATGCTCTGGAGACCTTGGGACCCCACTATTTTAAGCACTGGCGCTATGAACCAATTC 634
```

Db 655 AATGTGATGGAGCTCAAGGACAGATCTTTCTCCTGAAATCATGCCACTTTTGGCCCCAAG--- 711  
QY 635 AGCCCCCTACTGTCGGTCTTCGGCATTTGGGACCTCGTGGCCCAAGCTGGAGGACCTTC 694  
Db 712 AACCACTACTCCCCATCTTCGACTGGGCTCCATCGCTCGGTCGGGCGGAGCGACTTC 771  
QY 695 GAGGACCTGGGCTGCTGGGTGGCTCTGTAGGCATCAGATTCACTGGGATTTGTACCTG 754  
Db 772 CAGGATATAGCCCTCGAGGTGGCTGATAGGAATTAATATTGAATGAACTGTGATCTT 831  
QY 755 GACACCGGGGACTCTGGCTGTGGCTCCTCACTACTCTCTTCAGCTGCGAGGAGAGA----- 809  
Db 832 GATAAAGTGGCTCTGTAGTGGACCCCTCACTATCTTTTAAAGCTCTGGCAATAAACTT 891  
QY 810 -----GCTACAACTTCAGGACAGCCACTCACTGGTGGGAGCAACG 850  
Db 892 TCAGAAGTCTGCTCCTCGGGTACAACTTCAGATTGGCAGATATTACCAGAGCGAGCC 951  
QY 851 GGTTGGAGGCCCGACCTCTCAAGCTCTATGGAATTCGGCTTCGACATCTCTGTCACC 910  
Db 952 GGGGTGGAGTTCCGACCCCTGATGAAGCCCTACGGGATCCGCTTGACGTGATGGTGAAC 1011  
QY 911 GGGCAGGCGAGGAGTTCCG 930  
Db 1012 GGCAAGGTGCTTTCTTCTG 1031

## RESULT 6

US-10-342-887-786  
; Sequence 786, Application US/10342887  
; Publication No. US20040058340A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter S.  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Christopher J.  
; APPLICANT: Van de Veer, Laura Johanna  
; APPLICANT: Van de Vijver, Marc J.  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-188-999  
; CURRENT APPLICATION NUMBER: US/10/342,887  
; CURRENT FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: 60/298,918  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/380,710  
; PRIOR FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: 10/172,118  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 786  
; LENGTH: 1978  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-342-887-786

Query Match 9.0%; Score 243.6; DB 16; Length 1978;  
Best Local Similarity 58.3%; Pred. No. 3.9e-57;  
Matches 501; Conservative 0; Mismatches 329; Indels 30; Gaps 3;  
QY 98 TGGGCTCTCTCCGCAAAAAGGTATACAGAGCGGACCTGGAAACCCCAAGTTTTCATC 157  
Db 175 TGGGTGTTCTGATAAAGAGGGTTACCAAGAGCTCGACCTCTCCCTGAGAGTGTGTC 234  
QY 158 ATCCAAACTCAAGGGTTTCGTCAGTCAAGAGCTTGGAAACCGGCTGTGG 217  
Db 235 ATCCAAAGTCAAGGGCGTGGCTTCCCAACACCTCGGATCTTGGGAGCGGATCTGG 294  
QY 218 GATGTGGCGGACTTCTGTGAACCCACTCAGGAGAGAGAGCTTCTTCTTGGTGAACCAAC 277  
Db 295 GATGTGGCGGACTAGTCAITTCAGCCAGGAGAGAGAGCTTTTGTGGTCAACCAAC 354

QY 278 TTCTTGTGACCCAGCCCAAGTTTCAGGCGAGATGCCAGAGCAACCCCTCGTCCCTCCACTG 337  
Db 355 CTGATTGTGACCCCAACCCAGCGGAGAACTGCTGTGCTGAGATGAAGGCAATCTCTGAT 414  
QY 338 GCTAATCTGTGGTTCGAGAGGACTGCCCCGAAGGGAGGAGGACACACAGCCAGGCT 397  
Db 415 GCGCGGTGCTCCAAAGACAGCGACTGCCAGCTGGGGAAGCGGTTACAGCTTGGAAACGGA 474  
QY 398 GTAAAAACAGGCCAGCTGTGTGGTG---TTCAATGGGACCCACAGGACCTGTGAGATCTGG 454  
Db 475 GTGAGACCGGCGCTGCTCGGAGAGGGAATTTGGCCAGGGGACCTGTGAGATCTTT 534  
QY 455 AGTTGGTCCCAAGTGGAGAGTGGGCTGTGCGCTCGAGGCGCCCTGTGCGCCAGGCCAG 514  
Db 535 GCGGTGTCCTTGGAGACAAGCTCCAGGCGGAGGAGCAATTCCTGAAGGAGGCGGAA 594  
QY 515 AACTTCACACTGTTTCATCAAAACACAGTCACTTCAGCAAGTTCAACTTCTTAAGTCC 574  
Db 595 GACTTCACCAATTTTCATAAGAACCAATTCGTTTCCCAATTCACACTTCTCCAAAC 654  
QY 575 AATGCTTGGAGACCTGGGACCCCACTATTTTAAAGCACTCCGCTATGAACCAAAATTC 634  
Db 655 AATGTGATGGAGCTCAAGGACAGATCTTTCTGAAATCATGCCACTTTTGGCCCAAG--- 711  
QY 635 AGCCCTACTGTCCTGTTCCGATTTGGGACCTCTGGGCAAGGCTGGAGGACCTTC 694  
Db 712 AACCACTACTGCCCATCTTCGACTGGGCTCCATCTGCTGGGCGGAGGAGCACTTC 771  
QY 695 GAGGACCTGGGCTGCTGGTGGCTCTGTAGGCATCAGAGTTCACTGGGATTTGACCTG 754  
Db 772 CAGGATATAGCCCTCGAGGTGGCTGATAGGAATTAATATTGAATGAACTGTGATCTT 831  
QY 755 GACACCGGGGACTCTGGTGTGCTGGCTCACTACTCTCTTCAGCTGCGAGGAGAA----- 809  
Db 832 GATAAAGTGTCTGTAGTGGCACCCCTCACTATCTTTTGTAGCCGCTCTGGACAATAAACTT 891  
QY 810 -----GCTACAACTTCAGGACAGCCACTCACTGGTGGGAGCAACCG 850  
Db 892 TCAGAAGTCTGCTCCTCGGGTACAACTTCAGATTGGCAGATATTACCAGAGCGAGCC 951  
QY 851 GGTGTGGAGGCCCGACCCCTCTCAAGCTCTATGGAATTCGCTTCGACATCTCTGTCACC 910  
Db 952 GGGGTGGAGTTCCGACCCCTGATGAAGCCCTACGGGATCCGCTTGTGAGTGTGTAAC 1011  
QY 911 GGGCAGGCGAGGAGTTCCG 930  
Db 1012 GGCAAGGTGCTTTCTTCTG 1031

## RESULT 7

US-09-833-082-1  
; Sequence 1, Application US/09833082  
; Patent No. US20020151480A1  
; GENERAL INFORMATION:  
; APPLICANT: Chun, Miyoung  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE USING 10218  
; FILE REFERENCE: MNI-227  
; CURRENT APPLICATION NUMBER: US/09/833,082  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1389  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-09-833-082-1

Query Match 8.9%; Score 239.2; DB 9; Length 1389;  
Best Local Similarity 56.3%; Pred. No. 5.7e-56;  
Matches 529; Conservative 0; Mismatches 378; Indels 33; Gaps 3;  
QY 152 TCATCATCAACCAACTCAAGGGGTTTCCCTCACTCAGATCAAGGAGCTTGGAAACCGG 211

```
Db 214 TCCGTTACGACCAAGGTCAAGGGCGTGTGTGACCAACATTTAAACCTTGGATTCGG 273
Qy 212 CTGTGGGATGTGGCCCACTTCTGTGAAGCCACTCAGGAGAGAACTGTTCTTCTTGTG 271
Db 274 ATCTGGGATGTGGCGGATATGTGATACCGCTCAGGAGGAATCTCCCTCTTGTGATG 333
Qy 272 ACCAACTTCTTGTGACCGCAGCCCAAGTTTCTGGGAGATGCCAGACACCCGTCGTC 331
Db 334 ACCAACTGATCTCACCATCAACAGACACAGGCGCTGTGCCCGAGATTC---CAGAT 390
Qy 332 CCACCTGGTAACTGTGCTGACGAGGACTGCCCGAAGGGAGGAGGACACACAGC 391
Db 391 GGGACCATGTGTGTAATCATGATGCCAGCTGTACTGCCGCTGTGCCGACCCACAGC 450
Qy 392 CACGCTGTAAACAGGCGAGTGTGTGTTCATATGGACCCACAGACCTGTGAGATC 451
Db 451 AACGAGTCTCAACAGGAGGTGCTGTAGCTTTCACGCGTCCGTCAAGACGTGTGAGTG 510
Qy 452 TGGAGTTGGTCCAGTGGAGAGTGC---GTTGTGCCCTCGAGGCCCTCTGCCGCGAG 508
Db 511 GCGGCTGTGTCGCGGTGGAGATGACACACAGTCCACACCTGCTTTTAAAGGCT 570
Qy 509 GCCAGAACTTCACTGTTTCAATAAAACACAGTCACTTTCAGCAAGTTCAACTTCT 568
Db 571 GCAGAAACTTCACTCTTTTGGTTAAGAACACATCTGGTATCCCAATTTAATTCAGC 630
Qy 569 AAGTCCATGCTTGGACACCTGGACCCCACTTATTTTAAAGCTGCGGCTATGAACCA 628
Db 631 AAGAGGATATCTTCCCAACATCACCACTACTTACCTCAAGTGTGANTTATGATGCT 690
Qy 629 CAATTCAGCCCTACTGTCCTGCTTCCGATTTGGGACCTCGTGGCCAGAGTGGAGGG 688
Db 691 AAAACAGATCCCTTCTGCCCATATTCCTGTTTGGCAAAATAGTGGAGACGAGACAC 750
Qy 689 ACCTTCAGGACCTGTCGCTGTGGTGTCTGTAGGCATCAGAGTTCAGTGGATGT 748
Db 751 AGTTTCCAGACATGGCGCTGGAGGAGGCATCATGGGCATCCAGTCAACTGGGACTGC 810
Qy 749 GACCTGGACACGGGACTCTGGCTGTGGCTCACTACTCTCTTCCAGCTGCAGAGA-- 806
Db 811 AACCTGGACAGCGCCCTCCCTGCTTGGCTTCCAGGACTCTCTTCCGCGCTCGATACA 870
Qy 807 -----AGAGCTACAACTTTCAGACAGCCACTCACTGTTGG 841
Db 871 CGGAGCTTGAGCACACATGATCTCTGCTGCTACAAATTTTCAGTTTGGCAAGTACTACAGA 930
Qy 842 GAGCAACCGGTGTGGAGCCCGACCTGCTCAAGCTCTATGGAATCCGCTTCGACATC 901
Db 931 GACCTGGTGGCAACGACGACGACGCTCATCAAGGCTTATGGCATCCGCTTCGACATC 990
Qy 902 CTGCTCACGGGACGAGGGAAGTTCCGGCTCATCCCGCGCGTCCACACTGGGCACC 961
Db 991 ATTGTTTGGAGGACAGGAAATTTGACATCATCCCATATGATCAACATCGGCTCT 1050
Qy 962 GGGGAGCTTGGTGGCGGTGGTCACTTTTCTGTGACCTGCTACTGCTGTATGTGAT 1021
Db 1051 GGCCTGGCACTGCTAGGATGGCAGCGCTGTGTGTGACATCATAGTCTCTACTGCTATG 1110
Qy 1022 AGAGAGCCCATTTCTACTGGAGGACAAAGTATGAGGAGG 1061
Db 1111 AAGAAAGACTTACTATTCGGAGAGAAATATAAATATG 1150

RESULT 8
US-10-305-720-831
; Sequence 831, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
```

```
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 831
; LENGTH: 1750
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 555697
US-10-305-720-831
```

```
Query Match      8.9%; Score 239.2; DB 16; Length 1750;
Best Local Similarity 56.3%; Pred. No. 6.2e-56;
Matches 529; Conservative 0; Mismatches 378; Indels 33; Gaps 3;
```

```
Qy 152 TCCATCATCAACAACTCAAGAGGGTTTCCGTCACATCAGATCAAGAGCTTGGAAACCG 211
Db 214 TCCGTTACGACCAAGGTCAAGGCGTGTGTGACCAACATTTAACTTGGATTCGG 273
Qy 212 CTGTGGGATGTGGCCCACTTCTGTGAAGCCACTCAGGAGAGAACTGTTCTTCTTGGTG 271
Db 274 ATCTGGGATGTGGCGGATATGTGATACCACTCAGAGTCCAGAGGAAACTCCCTCTCGTCATG 333
Qy 272 ACCAACTTCTTGTGACGCGCAGCCCAAGTTTCTGGGAGATGCCAGACACCCGTCGTC 331
Db 334 ACCAACTGATCTCTACCATGAACACAGACACAGGCGCTGTGCCCGAGATTC---CAGAT 390
Qy 332 CCACCTGGCTAACTGCTGGGTCCAGGAGACTGCCCGAAGGGAGGAGGACACACAGC 391
Db 391 GCGACCACTGTGTGTAATCAGATGCCAGCTGTACTGCCGCTCTGCCGACCCACAGC 450
Qy 392 CACGCTGTAAACAGGCGAGTGTGTGTTCATATGGACCCACAGGACCTGTGAGATC 451
Db 451 AACGAGTCTCAACAGGAGGTGCTGTAGCTTTCACGCGTCCGTCAAGACGTGTGAGGTG 510
Qy 452 TGGAGTTGGTGGCCAGTGGAGAGTGC---GTTGTGCCCTCGAGGCCCTCTGTCGCCAG 508
Db 511 GCGGCTGTGTCGCGGTGGAGATGACACACAGTCCACAACTGCTTTTAAAGGCT 570
Qy 509 GCCAGAACTTCACTGTTTCAATAAAACACAGTCACTTTCAGCAAGTTCAACTTCTCT 568
Db 571 GCAGAAACTTCACTCTTTTGGTTAAGAACACATCTGGTATCCCAATTTAATTCAGC 630
Qy 569 AAGTCCATGCTTGGAGACCTGGGACCCCACTATTTTAAAGCTGCGCTATGAACCA 628
Db 631 AAGAGGATATCTTCCCAACATCACTACTTACCTCAAGTGTGANTTATGATGCT 690
Qy 629 CAATTCAGCCCTACTGTCCTGCTTCCGATTTGGGACCTCGTGGCCAGAGTGGAGGG 688
Db 691 AAAACAGATCCCTTCTGCCCATATTCCTGTTTGGCAAAATAGTGGAGACGAGACAC 750
Qy 689 ACCTTCAGGACCTGTCGCTGTGGTGTCTGTAGGCATCAGAGTTCAGTGGATGT 748
Db 751 AGTTTCCAGACATGGCGCTGGAGGAGGCATCATGGGCATCCAGTCAACTGGGACTGC 810
Qy 749 GACCTGGACACCGGAGACTCTGGCTGTGGCTCACTACTCTCTTCCAGCTGCAGAGA-- 806
Db 811 AACCTGGACAGCGCCCTCCCTGCTTGGCTTCCAGGACTCTCTTCCGCGCTCGATACA 870
Qy 807 -----AGAGCTACAACTTTCAGACAGCCACTCACTGTTGG 841
Db 871 CGGAGCTTGAGCACACATGATCTCTGCTGCTACAAATTTTCAGTTTGGCAAGTACTACAGA 930
Qy 842 GAGCAACCGGTGTGGAGCCCGACCTGCTCAAGCTCTATGGAATCCGCTTCGACATC 901
Db 931 GACCTGGTGGCAACGACGACGACGCTCATCAAGGCTTATGGCATCCGCTTCGACATC 990
Qy 902 CTGCTCACGGGACGAGGGAAGTTCCGGCTCATCCCGCGCGTCCACACTGGGCACC 961
Db 991 ATTGTTTGGAGGACAGGAAATTTGACATCATCCCATATGATCAACATCGGCTCT 1050
Qy 962 GGGGAGCTTGGTGGCGGTGGTCACTTTTCTGTGACCTGCTACTGCTGTATGTGAT 1021
Db 1051 GGCCTGGCACTGCTAGGATGGCAGCGCTGTGTGTGACATCATAGTCTCTACTGCTATG 1110
Qy 1022 AGAGAGCCCATTTCTACTGGAGGACAAAGTATGAGGAGG 1061
Db 1111 AAGAAAGACTTACTATTCGGAGAGAAATATAAATATG 1150
```



962 GGCGCAGCTTGGCTGGCGTGTCACTTTTCTGTGACCTGCTACTGCTGATGTGGAT 1021  
QY  
1051 GGCCTGGCACTGCTAGGCAATGGCGACGGTGTGTGACATCAATGCTCTACTGTCATG 1110  
Db  
1022 AGAAGCCCATTTCTACTGGAGGCAAAAGTATGAGGAGG 1061  
QY  
1111 AAGAAAGAGCTCTACTATCGGAGAAGAAATATATATATG 1150  
Db

## RESULT 9

```

US-10-128-558-20
; Sequence 20, Application US/10128550
; Publication No. US20040219521A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Zhiwei
; APPLICANT: Weng, Gezhi
; APPLICANT: Boyle, Bryan J
; APPLICANT: Dmanac, Radoje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 812A
; CURRENT APPLICATION NUMBER: US/10/128,558
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: US 60/339,453
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; Remaining Prior Application data removed - 8
; NUMBER OF SEQ ID NOS: 412
; SOFTWARE: ptl_FL_genes Version 6.0
; SEQ ID NO 20:
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1269)
US-10-128-558-20

```

332	QY	CCACTGGCTTA	CTGCTGGT	CGACAGAGACT	GC	CCCCGAAGGGAGGGAGGCACACACAGC	391
364	DB	GGGACCACT	GTGTGTAAT	CAATCAGAT	CC	CAGCTGTACTGCCGGCTCTGCGGCACCCACACAGC	423
392	QY	CACGGTGTAAAA	CAAGGCAGT	GTGTGTGTTT	CAAT	TGGGACCCACAGGACCTGTGAGATC	451
424	DB	AACGAGTCT	CAACAGCAG	GTGCGTAGCT	TTCA	ACGGGTCTGTCAAGACGTGTGAGGTG	483
452	QY	TGGAGTTGGT	GCCCACT	GTGAGAGTGGC	--	GTGTGTCCTCGAGGCCCTGCTGGCCACG	508
484	DB	GGCGCCT	GTGTCGCGGT	GGAGATGACA	CA	CACGTGCCACACCTGCTTTTTAAAGGCT	543
509	QY	GGCCAGAACTT	CACACTGT	TTCATCAAA	AACACAGT	CACCTTCAGCNAAGTTCAACTTCTCT	568
544	DB	GCAGAAACTT	CACCTCTTT	TGGTTAAGA	ACAACAT	CTGSTATCCAAATTTAAATTTACAGC	603
569	QY	AGTCCAA	TGCTTGGAG	ACCTGGGAC	CCGACCTAT	TTTAAAGCACTCGGCTATGAACA	628
604	DB	AAGAGGA	TATCCCTTCC	CAACATCA	CCCACTACTT	ACCTCAAGTGTGATTTATGATGCT	663
629	QY	CAATTCAG	CCCTACTGT	CTCCGCTGT	TCGCGAAT	TGGGACCTCGTGGCCCAAGCTGGAGGG	688
664	DB	AAACAGAT	CCCTCTCTG	CCCCCAT	ATTCG	CTTGGCAAAATAGTGAAGACGACAGAC	723
689	QY	ACCTTCGAG	ACCTGGCT	TGCTGGTGGCT	CTGTAG	GCATCAGAGTTCACTGGGATGT	748
724	DB	AGTTTCCAG	GATGGCCGT	GGAGGGAGCA	TCA	TGGGCATCCAGGTCAACTGGGACTGC	783
749	QY	GACCTGG	ACACCGGG	ACTCTGGCT	CTCACTACT	CTCTCCAGCTGCAGAGAGA--	806
784	DB	AACCTGAG	ACAGCGCCT	CCCTCTGCT	TGCCAGGTACT	CCTTCGCGCCTCGATACA	843
807	QY	-----	-----	-----	AGAGTCA	AACTTCAGACAGCACTCACTGGTGG	841
844	DB	CGGAGCT	TGAGACA	CAACGTAT	CTCTCGCTAC	AAATTTCCAGTTTGGCAAGTACTACAGA	903
842	QY	GAGCAAC	CGGTGTGG	AGCGCCG	CACCTGTCT	CAAGCTCTATGGAATCCGTTTCGACATC	901
904	DB	GACCTGGT	TGGCAAC	GACGACG	CA	CGTCAACAAGCCTATGSCATCCGCTTCGACATC	963
902	QY	CTCGTCA	CCGGCAGC	AGGAAGT	TGGGCTCAT	CCCAAGCGCGGTCACTGGGCAAC	961
964	DB	ATTGTGT	TTGGGAG	GCAGGAAAT	TTTGACAT	CATCCCCCACTATGATCAACATCGGCTCT	1023
962	QY	GGGGCAG	CTCGCT	TGGGGGT	GTGCACCT	TTTTTGTGTGACCTGTACTCTGTATGTGGAT	1021
1024	DB	GGCTGG	CATGCTAG	GATGGCA	TGGCACCCTG	CTGTGTGACATCATAGTCTCTACTGCA	1083
1022	QY	AGAGAAG	CCCCATTT	CTACTGG	AGGACA	AAAGTATGAGGAGG	1061
1084	DB	AAGAAAG	ACTCTACT	ATCTCGG	AGAAGAAAT	TATAAATATG	1123

```

RESULT 10
US-10-386-414-18
; Sequence 18, Application US/10386414
; Publication NO. US2004006016A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Robison, Keith E.
; APPLICANT: White, David
; APPLICANT: Williamson, Mark W.
; APPLICANT: Cook, William James
; APPLICANT: Meyers, Rachel E.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: NOVEL 27875, 22025, 27420, 17906, 16319,
; FILE OF INVENTION: 55092 AND 10218 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPIO3-0210NMIM
; CURRENT APPLICATION NUMBER: US/10/386,414
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 09/426,282

```



; PRIOR FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 09/668,266
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/330,970
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: 09/724,599
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/860,193
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 10/283,023
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 60/335,044
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 10/010,943
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/254,037
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/833,082
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1167
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-386-414-18

Query Match
Best Local Similarity 8.8%; Score 237.2; DB 16; Length 1167;
Matches 527; Conservative 1; Mismatches 379; Indels 33; Gaps 3;

QY 152 TCCATCATCAACAACTCAAGGGTTCCTCGTCTCAGATCAAGAGCTTGGAAACGG 211
Db 187 TCCGTTACGACCAAGTCAAGGGCTGTGTGACCAACATCTTAAACTTTGGATTCCGG 246
QY 212 CTGTGGGATGGCGGCTTCTGTGAAGCCACTCAGGAGAGAAAGTCTTCTTGGTG 271
Db 247 ATCTGGGATGGCGGATATGTATACCAAGTCTAGGAGGAAATCCCTCTTGTGATG 306
QY 272 ACCAACTTCTTGTGACGCCACCCCAAGTTCAGGGCAGATGCCAGAGACCCGTCCTGC 331
Db 307 ACCAAGGTGATCTTCAACATGAACACAGACAGAGGCTGTGCCCGAGATTC--CAGAT 363
QY 332 CCATCTGGCTTAAGTCTGGGTGACAGAGACTGCCCGGAGGGAGGAGGACACACAGC 391
Db 364 GCGACCACTGTGTGTAATCAGATCCAGCTGTACTGCGGGCTCTGCGGACCCACAGC 423
QY 392 CACGCTGTAAACACAGCCAGTGTGTGTGTTCAATGGACCCACAGGACCTGTGAGATC 451
Db 424 AACGAGTCTCAACAGGAGGTGCGTAGCTTCAACGGGTCTGTCAAGACGTGTGARGTG 483
QY 452 TGGAGTTGGTCCCAAGTGGAGAGTGGC---GTGTGCCCCCTCAGGCCCTCTGCGCCAG 508
Db 484 GCGGCTGTGTGCGGCTGGAGGATGACACACAGCTGCCACACCTGCTTTTAAAGGCT 543
QY 509 GCGGAGACTTCACTGTTTCAATCAAAACACAGTCACTTCAAGAGTTCAACTTCTCT 568
Db 544 GAGAAACTTCACTCTTTTGGTTAAGAACACATCTGGTATCCCAAAATTTAATTCAGC 603
QY 569 AAGTCCAATGCTTGGAGACCTGGGACCCCACTATTTTAAGCACTGCGGCTATCAACA 628
Db 604 AAGAGGATATCTTCCCAACATCACTACTTACCTCAAGTCGTGATTTATGATGCT 663
QY 629 CAATTACGCCCTTCTGTCGCTGTTCCGATTTGGGACCTCTGTGGCCCAAGCTGGAGG 688
Db 664 AATAAGATCCCTTCTGCCCCATATTCGCTTGTGCAAAATAGTGGAGACGACGAGCAC 723
QY 689 ACCTTCGAGACCTGGGCTGTGTGGTGTCTGTAGGCATCAGAGTTCACTGGGATGTT 748
Db 724 AGTTTCCAGACATGGCGGTGAGAGGAGGCATCATGGGCATCCAGTCACTGGGACTGC 783
QY 749 GACCTGGACACGGGAGTCTGGCTGTCTGGCTCACTACTCTTCCAGTCCAGGAGA-- 806

784 AACCTGGACAGAGCGGCTCCCTCTGCTTGGCCAGGACTCTCTTCCGCGCCTCGATACA 843
QY 807 -----AGAGCTACAACTTCAGGACAGCCTCACTCTGTTGG 841
Db 844 CGGAGCCTTGAGCACACAGTATCTCTGCTGCTACAAATTCAGGTTTGCCAAGTACTACAGA 903
QY 842 GAGCAACCGGGTGTGGAGGCCGACCCCTCTCAAGCTCTATGAAATCCGCTTTCGACATC 901
Db 904 GACTGTGCTGCAACGAGCAGCAGCCTCATCAAGCCTATGGCATCCGCTTTCGACATC 963
QY 902 CTCTACCGGCGGAGGAGGAGTTTGGGCTCATCCCCACGCGCTCACTCTGTTGGAT 1021
Db 964 ATTGTGTTGGGAAGGAGGAGGAAATTTGACATCATCCCACTATGATCAATCGGCTCT 1023
QY 962 GGGCAGCTTGGCTGGCGTGGTCACTTTTCTGTGACCTGCTGCTGTTGTTGTTGTTG 1021
Db 1024 GGCCTGCACTGTAGGCATGGGAGCGGTCTGTGACATCATAGTCTCTTACTGATG 1083
QY 1022 AGAAGAGCCATTTCTACTCGAGGACAAAGTATGAGGAGG 1061
Db 1084 AAGAAAGACTCTACTATCGGAGAGAGATATAAATATG 1123

RESULT 11
US-10-386-414-16
; Sequence 16, Application US/10386414
; Publication No. US2004006016A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Robison, Keith E.
; APPLICANT: White, David
; APPLICANT: Williamson, Mark W.
; APPLICANT: Cook, William James
; APPLICANT: Meyers, Rachel E.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: NOVEL 27875, 22025, 27420, 17906, 16319,
; FILE REFERENCE: MPI03-0210MNM
; CURRENT APPLICATION NUMBER: US/10/386,414
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 09/426,282
; PRIOR FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 09/668,266
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/330,970
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: 09/724,599
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/860,193
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 10/283,023
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 60/335,044
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 10/010,943
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/254,037
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/833,082
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 2048
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-386-414-16

Query Match
8.8%; Score 237.2; DB 16; Length 2048;
Best Local Similarity 56.1%; Pred. No. 2.4e-55;
Matches 527; Conservative 1; Mismatches 379; Indels 33; Gaps 3;



Db 1153 CCGGAGCTTTGAGACAAAGTATCTCTCTGGCTACAAATTCAGGTTTGGCAAGTACTACAGA 1212  
QY 842 GAGCAACCGGGTGTGAGAGCCCGACCTCTGCTCAAGCTCTATGGAATCGGTTTCGACATC 901  
Db 1213 GACCTGGCTGGCAACGAGCAGCGCAGCTCATCAAGGCTATGCGATCGCTTCGACATC 1272  
QY 902 CTCGTCACCGGAGCAGCGAGGGAAGTTCGGGCTCATCCCCACGCGCTCACACTGGGACC 961  
Db 1273 ATTGTGTTTGGGAAGCAGGGAATTTGACATCATCCCCACTATGATCAACATCGGCTCT 1332  
QY 962 GGGGAGCTTTGGCTGGGCGTGTGCACCTTTTCTGTGACCTGCTACTGCTGTATGTGGAT 1021  
Db 1333 GCGCTGGCACTGCTAGGATGCGACCGTGTGTGTGACATCATAGTCTCTACTGCAATG 1392  
QY 1022 AGAGAAGCCCATTTCTACTGAGAGCAAAAGTATGAGAGG 1061  
Db 1393 AAGAAAAGACTCTACTATCGGAGAGAAATATAAATATG 1432

## RESULT 13

US-10-187-659A-4  
; Sequence 4, Application US/10187659A  
; Publication No. US20040002152A1  
; GENERAL INFORMATION:  
; APPLICANT: Kenneth W. Doble  
; TITLE OF INVENTION: ANTISENSE MODULATION OF P2X4 EXPRESSION  
; FILE REFERENCE: RTS-0379  
; CURRENT APPLICATION NUMBER: US/10/187,659A  
; CURRENT FILING DATE: 2002-07-01  
; NUMBER OF SEQ ID NOS: 143  
; SEQ ID NO 4  
; LENGTH: 2048  
; TYPE: DNA  
; ORGANISM: H. sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (310) ... (1476)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 789  
; OTHER INFORMATION: n = A,T,C or G  
US-10-187-659A-4

Query Match 8.8%; Score 236.6; DB 16; Length 2048;

Best Local Similarity 56.1%; Pred. No. 3.5e-55;

Matches 527; Conservative 0; Mismatches 380; Indels 33; Gaps 3;

QY 152 TCCATCATCACCACCAACTCAAGGGGTTTCCGTCATCATGATCAAGGACCTGGAAACCGG 211  
Db 496 TCCGTTACGACCAAGGTCAAGGGGTGGTGTGACCAACACTTCTAAACITGGATTCGG 555  
QY 212 CTGTGGGATGTGGCGCACTTCGTGAAGCCACTCTAGGAGAGAACGTGTTCTTCTTGTG 271  
Db 556 ATCTGGGATGTGGCGGATTAATGATACAGCTCAGGAGAACTCCCTCTTGTGATG 615  
QY 272 ACCACTTCTTGTGAGCGGACCCCAAGTTTCAAGGAGATGCCAGACACCCCTCCGTC 331  
Db 616 ACCAAGTGTATCCTCACCATGAACACAGACAGGGGCTGTGCCGAGATTCTCAGAT 672  
QY 332 CCACCTGGCTACTGTGGTTCGACGAGCTCCCGGAGGGAGGAGGACACACAGC 391  
Db 673 GGGACACTGTGTAATATCATGATCCAGCTGTACTGCCGGCTGTGCCGGACCCACAGC 732  
QY 392 CACGGTGAATAAACAGGCGAGTGTGTGTTTCAATGGGACCCACAGACCTGTGAGATC 451  
Db 733 AACGGAGTCTCAACAGCAGGTGCGTAGCTTTCACGGGTCTGTCAAGACGTGTGANGTG 792  
QY 452 TGGAGTTGTGCCAGTGTGAGATGGC---GTTGTGCCCTGAGGCCCTGCTGCCCGAG 508  
Db 793 GCGGCTGTGTGCCGGTGTGAGATGACACACAGCTGCCACCACTGCTTTTAAAGGCT 852  
QY 509 GCCAGAACTTCACACTGTTCATCAAAAACACAGTCACCTTCAGCAAGTTCAACTTCCT 568

Db 853 GCAGAAAACITTCATCTTTTGGTTAAGAACAAACATCTGGTATCCCAATTTAATTTACG 912  
QY 569 AAGTCCAAATGCCTTGGAGACCTGGGACCCACCTATTTTAAAGCACTGCCGCTATGAACCA 628  
Db 913 AAGAGGAATATCTTCCCAACATCACCACTACTTACCTCAAGTCGTGCAITTTATGATGCT 972  
QY 629 CAATTCAAGCCCTTACTGTCCCGTGTTCGCAATTTGGGACCTCTGTGGCCAAAGGCTGAGGG 688  
Db 973 AAAACAGATCCCTTCTGCCCATATTTCCGCTTTGGCAAAATAGTGAGAAACGAGGACAC 1032  
QY 689 ACCTTCAGGACCTGGGCTTGTGGGTGGCTCTGTAGGCATCAGAGTTCACTGGGATTTGT 748  
Db 1033 AGTTTCCAGACATGCGCTGGAGGAGGATCATGGGATCCAGTCACTGGGACTGC 1092  
QY 749 GACCTGACACCGGGGACTCTGGCTGCTGGCTCTCACTTCTTCAGGTGACAGAGA-- 806  
Db 1093 AAGCTGGACAGAGCGGCTCCCTCTGCTTCCAGGTACTCTCTCCGCGCTCGATACA 1152  
QY 807 -----AGAGCTACAATTCAGGACACGCCACTCACTGCTGG 841  
Db 1153 CCGGACGTTGAGCACAAAGTATCTCTGGCTACAATTTAGGTTTGCAGTACTACAGA 1212  
QY 842 GAGCAACCGGGTGTGAGGCGCCGACCTCTCAAGCTCTATGGAATCCGCTTTCGACATC 901  
Db 1213 GACCTGGCTGCAACGAGCAGCGCTCATCAAGGCTATGCGATCCGCTTCGACATC 1272  
QY 902 CTGTCACCGGAGCAGGAGGAAAGTTCGGGCTCATCCGCGGCTGACACTGGGACCC 961  
Db 1273 ATTGTGTTTGGGAAAGCAGGGAATTTGACATCATCCCCACTATGATCAACATCGGCTCT 1332  
QY 962 GGGGAGCTTTGGCTGGGCGTGTCACTTTTCTGTGACCTGCTACTGCTGTATGTGGAT 1021  
Db 1333 GCGCTGGCACTGTAGGATGGCGCGTGTGTGTGACATCATAGTCTCTACTGCAATG 1392  
QY 1022 AGAGAAGCCCATTTCTACTGAGGACAAAGTATGAGAGG 1061  
Db 1393 AAGAAAAGACTCTACTATCGGAGAGAAATATAAATATG 1432

## RESULT 14

US-09-969-347-225  
; Sequence 225, Application US/09969347  
; Patent No. US20020115085A1

; GENERAL INFORMATION:

; APPLICANT: Ebner, Reinhard

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu

; FILE REFERENCE: Sets

; CURRENT APPLICATION NUMBER: US/09/969,347

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US/60/237,598

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: US/60/237,604

; PRIOR FILING DATE: 2000-10-03

; NUMBER OF SEQ ID NOS: 318

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 225

; LENGTH: 2633

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-969-347-225

Query Match 8.7%; Score 233.4; DB 9; Length 2633;  
Best Local Similarity 56.7%; Pred. No. 3e-54;  
Matches 526; Conservative 0; Mismatches 371; Indels 30; Gaps 4;

QY 164 AAATCAAGGGGTTTCCGTCATCATGATCAAGGACCTTGGAAACCGGCTGTGGGATG 223  
Db 398 AAATCAAGGGGCTTGGCGGTGACCCAGTCCCTGGGCTCGGCCCGGAGGTCTGGGATG 457  
QY 224 GCGGACTTCTGGAAGCCACTCAGGAGAGAAAGTCTTCTTGTGTGACCACTTCCTT 283  
Db 458 GCTGACTACGTTCTCCAGCCGAGGGGAGACAACTCTCTCGTGTGATGACCAATTCATC 517

QY 284 GTGACGCCAGCCAGTTTCAGGGCAGATGCCAGAGCACCGTCCGTCACCTGGCTAAC 343  
 DB GTGACCCCGAGCAGACTCAGGCTACTGCGCAGAGCACCC-----AGAGGGGGGATA 571  
 QY 344 TGCTGGGTGACGAGGACTGCCCGAAGAGGGAGGAGGACACACAGCCACGGTGTAAAA 403  
 DB TGAAGGAAGACAGTGGCTGTACCCCTGGGAAGGCCAAGAGAGGCCCAAGGATCCGC 631  
 QY 404 ACAGCCAGTGTGGTGTCAATGGGACCCACAGGACCTGTGAGACTGGAGTTGTGC 463  
 DB ACGGCAAGTGTGGCTTCAAGACACTGTGAAGACGTGTGAGATCTTTGGCTGTGTGC 691  
 QY 464 CCAGTGGAGAGTGGCGTTGTGCCCTCGAGGC---CCCTGTGGCCCGAGGCCAGAACTTC 520  
 DB CCCGTGGAGGTGGATGACGACATCCCGCGCCCTCTCTCGAGAGGCCGAGAACTTC 751  
 QY 521 ACAGTGTTCATCAAAAAACAGTACCTTACGAGTCAACTTCTTAAGTCCATGCC 580  
 DB ACTCTTTTCATCAAGAACAGATCAGCTTTCCACGCTTCAAGGTCAACAGCGCGCAACTG 811  
 QY 581 TTGGAGACCTGGAGCCCGCCACTATTTTAAAGCACTGCCGTATGAACACAAATTCAGCCCC 640  
 DB GTGAGGAGGTGAATGTGCCCAACATGAGACCTGCTCTTTCACAGACCTTCACCCC 871  
 QY 641 TACTGTCCCGTTCGCGCATTTGGGACCTGTGTGSCCAAGCTGTGAGGAGCCTTCGAGGAC 700  
 DB CTGTGCCCGAGCTTCCAGCTTGGCTACGTGTGCAAGAGTCAAGGCCAGAACTTCAGCACC 931  
 QY 701 CTGGGCTGTGGTGGCTGTAGGACATCAGTTCAGTGGGATTCGACCTGGACACC 760  
 DB CTGGCTGAGAGGGTGGAGTGGTGGCATCAGTGGCTCTTTCACAGACCTTCGACCTGG 991  
 QY 761 GGGGACTGTGGCTGTGGCTTCACTACTCTTCCAGCTGCAGGAGAGA----- 809  
 DB CAGGTACGGCACTGCAGACCCATCTATGAGTTCATGGCTGTACGAGAGAAAAATCTC 1051  
 QY 810 -----GCTACACTTCAGACAGCCACTCAGTGGTGGGAGCAACCGGTGTGGAGGCC 862  
 DB TCCCAGGCTTCAACTTCAGGTTTCCAGGACCTTGTGGAGAAC---GGACCACTAC 1108  
 QY 863 CGCACCTGTCTCAAGCTCTATGGAATCCGCTTTCACATCTCTCGTCAACCGGCGCAGGAGG 922  
 DB CGTCACTCTTCAAGGTGTGGGATTCGCTTTCACATCTCTGTGGAGCGCAAGGCCGG 1168  
 QY 923 AAGTTCGGGCTATCCCAAGCGGTTCACACTGGGACCGGGGAGCTTGGCTGGGCGTG 982  
 DB AAGTTTGACATCATCCCTACAAATGACCACCATCGGCTCTGGAATGGCATCTTTGGGGTG 1228  
 QY 983 GTCACCTTTTCTGTGACTCTGCTACTGTGTATGTGGATAGAGAGCCATTTCTACTGG 1042  
 DB GCCACAGTTCTGTGACTCTGCTCTGCTTCACTCTGCTTACATCTGCTTAGAGGCCTACTACA 1288  
 QY 1043 AGGACAAAGTATGAGGAGGCCCAAGGCC 1069  
 DB CAGAAGAGTTCAAATACGCTGAGGAC 1315

## RESULT 15

US-10-352-684A-53  
 ; Sequence 53, Application US/10352684A  
 ; Publication No. US20030215452A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Millennium Pharmaceuticals Inc.  
 ; APPLICANT: Carroll, Joseph M.  
 ; APPLICANT: Healy, Aileen  
 ; APPLICANT: Welch, Nadine S.  
 ; APPLICANT: Kelly, Louise M.  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
 ; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,  
 ; TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 1212,  
 ; TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,  
 ; TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES

; FILE REFERENCE: MPI02-019P1RNMNIM  
 ; CURRENT APPLICATION NUMBER: US/10/352,684A  
 ; CURRENT FILING DATE: 2003-01-28  
 ; PRIOR APPLICATION NUMBER: US 60/354,333  
 ; PRIOR FILING DATE: 2002-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/360,258  
 ; PRIOR FILING DATE: 2002-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/364,476  
 ; PRIOR FILING DATE: 2002-03-15  
 ; PRIOR APPLICATION NUMBER: US 60/375,626  
 ; PRIOR FILING DATE: 2002-04-26  
 ; PRIOR APPLICATION NUMBER: US 60/386,494  
 ; PRIOR FILING DATE: 2002-06-06  
 ; PRIOR APPLICATION NUMBER: US 60/390,965  
 ; PRIOR FILING DATE: 2002-06-24  
 ; PRIOR APPLICATION NUMBER: US 60/392,480  
 ; PRIOR FILING DATE: 2002-06-28  
 ; PRIOR APPLICATION NUMBER: US 60/394,128  
 ; PRIOR FILING DATE: 2002-07-03  
 ; PRIOR APPLICATION NUMBER: US 60/399,783  
 ; PRIOR FILING DATE: 2002-07-31  
 ; PRIOR APPLICATION NUMBER: US 60/403,221  
 ; PRIOR FILING DATE: 2002-08-13  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 62  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 53  
 ; LENGTH: 2643  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (174)...(1373)  
 ; US-10-352-684A-53

Query Match 8.7%; Score 233.4; DB 15; Length 2643;  
 Best Local Similarity 56.7%; Pred. No. 3e-54;  
 Matches 596; Conservative 0; Mismatches 371; Indels 30; Gaps 4;

QY 164 AAATCAAGGGGTTTCGTACTCAGATCAAGAGCTTGAACCGCTCTGGGATGTG 223  
 DB 375 AAATCAAGGGGCTCGCGGTGACCCAGCTCCTCGGCCCGCCAGGCTGGGATGTG 434  
 QY 224 GCGGACTTCTGTAAGCCACCTCAGGGAGAGAACTGTTCTTCTTGGTACCAACTTCTT 283  
 DB 435 GCTGACTAGCTTCTCCAGCCCGAGGGGAGCACTCTTCTGTTGTCATGACCAATTCATC 494  
 QY 284 GTGACGCCAGCCAGTTCAGGGCAGATGCCAGAGCACCGCTCGTCCGCTACCTGGCTAAC 343  
 DB 495 GTGACCCCGAAGCAGACTCAAGGCTACTGCGCAGAGCACCC-----AGAGGGGGGATA 548  
 QY 344 TGCTGGTGCAGAGAGTGTGCCCTCGAGGC---CCCTGTGGCCCGAGGCCAGAACTTC 403  
 DB 549 TGAAGGAAGACAGTGGCTGTACCCCTGGAGAGGCCAGAGAGGCCCAAGGATCCGC 608  
 QY 404 ACAGGCCAGTGTGTGGTGTCAATGGGACCCACAGGACCTGTGAGATCTGAGTGTGTGC 463  
 DB 609 ACGGCAAGTGTGTGGCTTCAACGACACTGTGAAGACGTGTGAGATCTTTGGCTGTGTGC 668  
 QY 464 CCAGTGGAGAGTGGGTGTGGCTTCGAGGC---CCCTGTGGCCCGAGGCCAGAACTTC 520  
 DB 669 CCCGTGGAGGTGGATGACGACATCCCGCGCCCTCTCTCGAGAGGCCGAGAACTTC 728  
 QY 521 ACAGTGTTCATCAAAAAACAGTCACTTCAAGAGTTCAACTTCTTAAGTCCAAATGCC 580  
 DB 729 ACTCTTTTCATCAAGAACAGATCAGCTTTCACGCTTCAAGGTCAACAGCGCAACTG 788  
 QY 581 TTGGAGACCTGGGAGCCCGCACTTATTTAAGCACTGCCCTATGAACCACTTTCAGCCCC 640  
 DB 789 GTGAGGAGGTGAATGCTGTCGCCACATGAAGACCTTGCCTCTTTCACAGACCTTCACCCC 848  
 QY 641 TACTGTCCCGTGTTCGCGCATTTGGGACCTGTGTGSCCAAGGCTGTGAGGAGCCTTCGAGGAC 700

```
Db 849 CTGTGCCAGTCTTCCAGCTTGGCTACGTGTGCAAGAGTCAGGCCAGAACTTCAGCACC 908
QY 701 CTGGCTTTCCTGGGTGGCTCTGTAGGCATCAGAGTTCACTGGGATTGTGACCTGGACACC 760
Db 909 CTGGCTGAGAGGGTGGAGTGGTTGGCATCACCATCGACTGGCACTGTGACCTGGACTGG 968
QY 761 GGGGACTCTGGCTGTGGCTCCTACTCTCTTCAGCTTCCAGCTGCGAGGAGA----- 809
Db 969 CACGTACGGCACTGCAGACCCCATCTATGAGTTCCATGGGCTGTACGAGAGAAAAATCTC 1028
QY 810 -----GCTACAACCTTCAGACACGCCACTCACTGGTGGGAGCAACCGGGTGTGGAGGCC 862
Db 1029 TCCCAGGCTTCAACTTCAGTTTGCCAGGCACITTTGGAGAC---GGGACCACTAC 1085
QY 863 CGCACCTGCTCAAGCTCTATGGAAATCCGCTTCGACATCCTCGTCAACCGGCGCAGGCAGGG 922
Db 1086 CGTCACCTCTTCAAGGTGTTGGGATTGGCTTTGACATCCTGGTGGACGGCAAGGCCGGG 1145
QY 923 AAGTTCGGGCTCATCCCGCCGCGTCACTGGGCGCGGCGAGCTTGGCTGGGGTG 982
Db 1146 AAGTTTGACATCATCCCTACAAATGACCACCATCGGCTCTGGAATTGGCACTTTGGGGTG 1205
QY 983 GTCACTTTTCTGTGACCTGCTACTGTGTATGTGGATAGAGAAGCCCATTTCTACTGG 1042
Db 1206 GCCACAGTTCCTGTGACCTGCTGCTGCTTCACATCCTGCTTACAGGCACTACTACAAG 1265
QY 1043 AGGACAAAGTATGAGGAGGCCAAGGCC 1069
Db 1266 CAGAGAAGTTCAAAATACGCTGAGGAC 1292
```

Search completed: November 22, 2004, 08:18:16  
Job time : 1337.37 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 224.653 Seconds  
(without alignments)  
8520.488 Million cell updates/sec

Title: US-09-820-095b-1  
Perfect score: 2693  
Sequence: 1 ttgtcgtactcgtgcccgc.....aaaaaaaaaaaaaaaaaaaaa 2693

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1515.4	56.3	1697	3	US-09-381-681-2
2	1162.4	43.2	1360	3	US-09-191-136-30
3	1155.8	42.9	1293	3	US-09-381-681-1
4	394.4	14.6	396	3	US-09-191-136-28
5	239.2	8.9	1750	4	US-09-016-434-831
6	239.2	8.9	1762	2	US-08-742-621-2
7	237.6	8.8	1206	3	US-09-191-608-21
8	233.4	8.7	2643	2	US-08-750-134A-10
9	233.4	8.7	2643	3	US-09-363-745-10
10	233.4	8.7	2643	4	US-09-023-655-897
11	225.4	8.4	1997	2	US-08-750-134A-6
12	225.4	8.4	1997	3	US-09-363-745-6
13	214.4	8.0	1421	3	US-09-191-608-14
14	213.2	7.9	1436	3	US-09-191-608-13
15	206	7.6	1837	2	US-08-750-134A-4
16	206	7.6	1837	3	US-09-363-745-4
17	200.2	7.4	237	3	US-09-191-136-29
18	199.4	7.4	1499	3	US-09-191-608-16
19	168.6	6.3	1243	3	US-09-191-136-15
20	156.2	5.8	1272	3	US-09-191-136-13
21	155.4	5.8	1349	3	US-09-191-608-15
22	155.4	5.8	1753	2	US-08-750-134A-8
23	155.4	5.8	1753	3	US-09-363-745-8
24	131.2	4.9	961	4	US-08-023-655-370
25	103.8	3.9	1853	3	US-08-842-079-19
26	103.8	3.9	1853	4	US-09-638-857-19
27	99.8	3.7	3540	3	US-08-842-079-16

28	99.8	3.7	3540	4	US-09-638-857-16	Sequence 16, Appl
29	94	3.5	94	3	US-09-191-136-18	Sequence 18, Appl
30	92.4	3.4	394	3	US-09-191-136-27	Sequence 27, Appl
31	90.4	3.4	878	1	US-07-915-934-3	Sequence 3, Appl
32	90.4	3.4	878	1	US-08-325-743-3	Sequence 3, Appl
33	83.2	3.1	531	3	US-09-191-608-8	Sequence 8, Appl
34	64.2	2.4	221	3	US-09-191-608-1	Sequence 1, Appl
35	64.2	2.4	910	3	US-09-191-608-2	Sequence 2, Appl
36	56.6	2.1	1872	3	US-09-801-052-1	Sequence 1, Appl
37	56.6	2.1	1872	4	US-10-020-121-1	Sequence 1, Appl
38	55.8	2.1	7218	4	US-08-232-463-14	Sequence 14, Appl
39	54.8	2.0	903	5	PCT-US95-06406A-21	Sequence 21, Appl
40	54.4	2.0	558	4	US-09-043-861-3	Sequence 3, Appl
41	54.2	2.0	183	3	US-09-014-416-60	Sequence 60, Appl
42	54.2	2.0	200	3	US-09-014-416-64	Sequence 64, Appl
43	54.2	2.0	9595	3	US-09-014-416-4	Sequence 4, Appl
44	54.2	2.0	9599	3	US-09-014-416-2	Sequence 2, Appl
45	54.2	2.0	9599	3	US-09-014-416-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1  
US-09-381-681-2  
; Sequence 2, Application US/09381681  
; Patent No. 6255472  
; GENERAL INFORMATION:  
; APPLICANT: TAKINO, Takashi  
; APPLICANT: NAKAMURA, Yusuke  
; TITLE OF INVENTION: HUMAN GENES  
; FILE REFERENCE: 055876  
; CURRENT APPLICATION NUMBER: US/09/381,681  
; CURRENT FILING DATE: 2000-01-10  
; EARLIER APPLICATION NUMBER: JPA 9-093044  
; EARLIER FILING DATE: 1997-03-26  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1697  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (46)..(1338)  
US-09-381-681-2

Query Match	56.3%	Score 1515.4	DB 3	Length 1697
Best Local Similarity	99.6%	Pred. No. 0		
Matches 1512	Conservative	0	Mismatches 6	Indels 0
				Gaps 0
QY	97	GTGGGCTCTCTCGCCAAAAAGGCTACAGAGCGGGACCTGGAAACCCAGTTTCCAT	156	
Db	180	GTGGGCTCTCTCGCCAAAAAGGCTACAGAGCGGGACCTGGAAACCCAGTTTCCAT	239	
QY	157	CATCAACAACTCAAGGGGTTTCCGTCTACTCAGATCAAGAGCTTGGAAACCGGCTGTG	216	
Db	240	CATCAACAACTCAAGGGGTTTCCGTCTACTCAGATCAAGAGCTTGGAAACCGGCTGTG	299	
QY	217	GGATGGCGGCTTCTGTAGCCCTCAGGAGGAGCGTGTCTTCTGTGACCAA	276	
Db	300	GGATGGCGGCTTCTGTAGCCCTCAGGAGGAGCGTGTCTTCTGTGACCAA	359	
QY	277	CTTCTTGTGACGCCACCCAAAGTTTCAGGCGAGATGCCAGAGCACCCCTCCGTCCTCACT	336	
Db	360	CTTCTTGTGACGCCACCCAAAGTTTCAGGCGAGATGCCAGAGCACCCCTCCGTCCTCACT	419	
QY	337	GGCTAACTCTGGGTTCAGAGACTGCCCCAAGGGGAGGAGGACACACAGCCACGG	396	
Db	420	GGCTAACTCTGGGTTCAGAGACTGCCCCAAGGGGAGGAGGACACACAGCCACGG	479	
QY	397	TGTAAAAACAGGCGAGTGTGTGTCTTCAATGGGACCCACAGACCTGTGAGATCTGGAG	456	

Db	480	TGTAATAAACAGGCCAGTGTGTGGTGTTCATTAATGGACCCACAGACCTGTGAGATCTGGAG	539
QY	457	TTGTGTGCCAGTGAAGATGGCGTTGTGCCCTGGAGCCCTCTGCTGGCCAGGCCAGAA	516
Db	540	TTGTGTGCCCGTGGAGAGTGGCGTTGTGCCCTGGAGCCCTCTGCTGGCCAGGCCAGAA	599
QY	517	CTTCACACTGTTTCATCAAAACACAGTCACCTTCAGCAAGTTCAACTTCTCTAACTTCT	576
Db	600	CTTCACACTGTTTCATCAAAACACAGTCACCTTCAGCAAGTTCAACTTCTCTAACTTCT	659
QY	577	TGCTTTGGAGACTGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAG	636
Db	660	TGCTTTGGAGACTGGACCCACCTATTTTAAGCACTGCGCTATGAACCAATTCAG	719
QY	637	CCCTACTGTCCTGTTCCGATGTTGGGACCTCGTGGCCAAAGGCTGGAGGACCTTCGA	696
Db	720	CCCTACTGTCCTGTTCCGATGTTGGGACCTCGTGGCCAAAGGCTGGAGGACCTTCGA	779
QY	697	GGACCTGGCGTTGTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTTGACCTGGA	756
Db	780	GGACCTGGCGTTGTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTTGACCTGGA	839
QY	757	CACCGGGACTCTGGCTGGCTCTACTCTCTCCAGTCGAGGAGAGCTACAA	816
Db	840	CACCGGGACTCTGGCTGGCTCTACTCTCTCCAGTCGAGGAGAGCTACAA	899
QY	817	CTTCAGGACAGCCACTCACTGGTGGGAGCAACCGGTTGTGGAGCCCGCACCTGTCTAA	876
Db	900	CTTCAGGACAGCCACTCACTGGTGGGAGCAACCGGTTGTGGAGCCCGCACCTGTCTAA	959
QY	877	GCTCTATGGAATCGCTTCGACATCTCTGTCAACGGGACGAGGAGTTCGGGCTCAT	936
Db	960	GCTCTATGGAATCGCTTCGACATCTCTGTCAACGGGACGAGGAGTTCGGGCTCAT	1019
QY	937	CCCAAGCGCTCACTATGGGACCGGGGACGCTGTGGCTGGCTGCTACCTTTTCTG	996
Db	1020	CCCAAGCGCTCACTATGGGACCGGGGACGCTGTGGCTGGCTGCTACCTTTTCTG	1079
QY	997	TGACCTGTACTGTGTATGTGATGAGAGCCATTTCTACTGGAGGACAAAGTATGA	1056
Db	1080	TGACCTGTACTGTGTATGTGATGAGAGCCATTTCTACTGGAGGACAAAGTATGA	1139
QY	1057	GGAGCCAAAGCCCGAAAGCAACCGCAACTCTGTGTGGAGGAGCTGCCCTTGTGATC	1116
Db	1140	GGAGCCAAAGCCCGAAAGCAACCGCAACTCTGTGTGGAGGAGCTGCCCTTGTGATC	1199
QY	1117	CCAAGCCGACTGGCCGAGTGCCTCAGACGAGCTCAGACCTGCAACCCACGCGCACTGC	1176
Db	1200	CCAAGCCGACTGGCCGAGTGCCTCAGACGAGCTCAGACCTGCAACCCACGCGCACTGC	1259
QY	1177	TGCTGGGAGTCAGACACAGACAGGATGCGCTGTCCAGTTCTGCACACCCACTTGC	1236
Db	1260	TGCTGGGAGTCAGACACAGACAGGATGCGCTGTCCAGTTCTGCACACCCACTTGC	1319
QY	1237	AACCAATTCGGGAGCTGTAGCCGTTCCCTGTCTGTGAGAGTTGGGGCTGGAGAGG	1296
Db	1320	AACCAATTCGGGAGCTGTAGCCGTTCCCTGTCTGTGAGAGTTGGGGCTGGAGAGG	1379
QY	1297	CGGGCCCTGCTGGGATCTCAGAGATGAGCCCGCAGATGGAGATTGGGGTAGAAT	1356
Db	1380	CGGGCCCTGCTGGGATCTCAGAGATGAGCCCGCAGATGGAGATTGGGGTAGAAT	1439
QY	1357	TCCACCTTTGAACCCACAGACAGTCCCTCCCTGTACTCCCACTTGTAGGGTGTGC	1416
Db	1440	TCCACCTTTGAACCCACAGACAGTCCCTCCCTGTACTCCCACTTGTAGGGTGTGC	1499
QY	1417	CTCAGGAGCCATGAAGTGTGGCTGTGTTTGTAGACGCGGACAACTGACCCGTGGAG	1476
Db	1500	CTCAGGAGCCATGAAGTGTGGCTGTGTTTGTAGACGCGGACAACTGACCCGTGGAG	1559
QY	1477	ACTGGGAGCCAGCAGGACCTGTATTGACGGGCTCCGACTGCATGTGGCAGGGGCTC	1536
Db	1560	ACTGGGAGCCAGCAGGACCTGTATTGACGGGCTCCGACTGCATGTGGCAGGGGCTC	1619

QY

1537

CTGCTGCGTCTGGGCTGGAGTCTCTCTCCAGTGTCTGTCTCCAGTGTCTCTAGCAG

1596

Db

1620

CTGCTGCGTCTGGGCTGGAGTCTCTCTCCAGTGTCTGTCTCCAGTGTCTCTAGCAG

1679

QY

1597

AGGTATGCTTACCAGCTG

1614

Db

1680

AGGTATGCTTACCAGCTG

1697

RESULT 2

US-09-191-136-30

; Sequence 30, Application US/09191136B

; Patent No. 6214581

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Lynch, Kevin J.

; APPLICANT: Burgard, Edward C.

; APPLICANT: Van Biesen, T.

; TITLE OF INVENTION: Nucleic Acids Encoding A Functional

; TITLE OF INVENTION: Human Purinoreceptor P2X3 and P2X6 And Methods Of Production

; TITLE OF INVENTION: And Use Thereof

; FILE REFERENCE: 6293.US.P1

; CURRENT APPLICATION NUMBER: US/09/191,136B

; CURRENT FILING DATE: 1998-11-13

; EARLIER APPLICATION NUMBER: US 09/008,526

; EARLIER FILING DATE: 1998-01-16

; EARLIER APPLICATION NUMBER: US 09/008,185

; EARLIER FILING DATE: 1998-01-16

; EARLIER APPLICATION NUMBER: US 60/071,298

; EARLIER FILING DATE: 1998-01-16

; EARLIER APPLICATION NUMBER: US 60/071,669

; EARLIER FILING DATE: 1998-01-16

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 30

; LENGTH: 1360

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Sequencing Primer (polynucleotide)

US-09-191-136-30

Query Match

43.2%

Score 1162.4;

DB 3;

Length 1360;

Best Local Similarity

99.9%;

Pred. No. 1.1e-288;

Matches 1163;

Conservative

0;

Mismatches

1;

Indels

0;

Gaps

0;

QY

97

GTGGCTCTCTCGCCAAAGGCTACAGGAGCGGACCTGGAAACCCAGTTTCCAT

156

Db

180

GTGGCGCTCTCTCGCCAAAGGCTACAGGAGCGGACCTGGAAACCCAGTTTCCAT

239

QY

157

CATCAACCAACTCAAGGGGTTTCGTCACATCAAGAGCTTGAACCCGCTGTG

216

Db

240

CATCAACCAACTCAAGGGGTTTCGTCACATCAAGAGCTTGAACCCGCTGTG

299

QY

217

GGATGGCCGACTTCGTGAAGCCACCTCAGGAGAGAACGTTTCTTCTTGGTGACAA

276

Db

300

GGATGGCCGACTTCGTGAAGCCACCTCAGGAGAGAACGTTTCTTCTTGGTGACAA

359

QY

277

CTTCTCTGTGAGCCAGCCCAAGTTTCAGGCGAGATGCCAGAGACCCGCTCCCTACT

336

Db

360

CTTCTCTGTGAGCCAGCCCAAGTTTCAGGCGAGATGCCAGAGACCCGCTCCCTACT

419

QY

337

GGCTAACTGCTGGGTGAGAGGACTGCCCGAAGGGAGGAGGACACACAGCCAGG

396

Db

420

GGCTAACTGCTGGGTGAGAGGACTGCCCGAAGGGAGGAGGACACACAGCCAGG

479

QY

397

TGTAAACAGCCGAGTGTGTCTTCAATGGGACCCACAGGACCTGTGAGATCTGAG

456

Db

480

TGTAAACAGCCGAGTGTGTGTCTTCAATGGGACCCACAGGACCTGTGAGATCTGAG

539

QY

457

TTGGTGTCCGAGTGGGCTGTGTGCTCTCGAGGCCCTGTGTGGCCCGAGGCCAGAA

516





Db 1095 GGAGGCCAAGGCCCGGAAAGCAACCGCAACTCTGTGTGAGGAGGTGGCCCTTGCATC 1154  
QY 1117 CCAAGCCGAGTGGCCGAGTGCCTTCAGACGAGAGTTCAGACCTTCGACCCACGCGCACTGC 1176  
Db 1155 CCAAGCCGAGTGGCCGAGTGCCTTCAGACGAGAGTTCAGACCTTCGACCCACGCGCACTGC 1214  
QY 1177 TGCTGGGAGTCAGACACGAGACACGAGATGGCCCTGTCCAAAGTCTTCGACACCCACTGGC 1236  
Db 1215 TGCTGGGAGTCAGACACGAGACACGAGATGGCCCTGTCCAAAGTCTTCGACACCCACTGGC 1274  
QY 1237 AACCCATTCCGGGAGCCTG 1255  
Db 1275 AACCCATTCCGGGAGCCTG 1293

RESULT 4  
US-09-191-136-28  
; Sequence 28, Application US/09191136B  
; Patent No. 6214581  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Lynch, Kevin J.  
; APPLICANT: Burgard, Edward C.  
; APPLICANT: Van Biesen, T.  
; TITLE OF INVENTION: Nucleic Acids Encoding A Functional  
; TITLE OF INVENTION: Human Purinoreceptor P2X3 and P2X6 And Methods Of Production  
; TITLE OF INVENTION: And Use Thereof  
; FILE REFERENCE: 6293.US.F1  
; CURRENT APPLICATION NUMBER: US/09/191.136B  
; CURRENT FILING DATE: 1998-11-13  
; EARLIER APPLICATION NUMBER: US 09/008.526  
; EARLIER FILING DATE: 1998-01-16  
; EARLIER APPLICATION NUMBER: US 09/008.185  
; EARLIER FILING DATE: 1998-01-16  
; EARLIER APPLICATION NUMBER: US 60/071.298  
; EARLIER FILING DATE: 1998-01-16  
; EARLIER APPLICATION NUMBER: US 60/071.669  
; EARLIER FILING DATE: 1998-01-16  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 28  
; LENGTH: 396  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Sequencing Primer  
US-09-191-136-28

Query Match 14.6%; Score 394.4; DB 3; Length 396;  
Best Local Similarity 99.7%; Pred. No. 1.2e-91;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 886 AATCGCTTCGACATCTCTGTCACCGGAGGAGGAGTGGGCTCATCCCGCAGCG 945  
Db 1 AATCGCTTCGACATCTCTGTCACCGGAGGAGGAGTGGGCTCATCCCGCAGCG 60  
QY 946 CGTCACACTGGGCAACCGGGGAGTGTGGGCTGAGTACCTTTTCTGTGACCTGCT 1005  
Db 61 CGTCACACTGGGCAACCGGGGAGTGTGGGCTGAGTACCTTTTCTGTGACCTGCT 120  
QY 1006 ACTGCTGTATGTGATGAGAGCCCACTTCTGAGGAGCAAGTATGAGGAGGCA 1065  
Db 121 ACTGCTGTATGTGATGAGAGCCCACTTCTGAGGAGCAAGTATGAGGAGGCA 180  
QY 1066 GGCCCCGAAGCAACCGGCAACTCTGTGTGAGGAGCTGGCCCTTGCATCCCAAGCCG 1125  
Db 181 GGCCCCGAAGCAACCGGCAACTCTGTGTGAGGAGCTGGCCCTTGCATCCCAAGCCG 240  
QY 1126 ACTGCCCCAGTGCCTCAGACGAGTTCAGACCTTCGACCCACGCGCACTGCTGGAG 1185  
Db 241 ACTGCCCCAGTGCCTCAGACGAGTTCAGACCTTCGACCCACGCGCACTGCTGGAG 300  
QY 1186 TCAGACACAGACACGAGGATGGCCCTGTCCAAAGTCTTCGACACCCACTTGCACACCCATTC 1245

Db 301 TCAGACACAGACACGAGATGGCCCTGTCCAAAGTCTTCGACACCCACTTGCACCCATTC 360  
QY 1246 CGGGAGCCTGTAGCGGTTCCTCTGTGTGAGATT 1281  
Db 361 CGGGAGCCTGTAGCGGTTCCTCTGTGTGAGATT 396

RESULT 5  
US-09-016-434-831  
; Sequence 831, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016.434  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 831:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1750 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: SCORNO101  
; CLONE: 555697  
US-09-016-434-831

Query Match 8.9%; Score 239.2; DB 4; Length 1750;  
Best Local Similarity 56.3%; Pred. No. 1.8e-51;  
Matches 529; Conservative 0; Mismatches 378; Indels 33; Gaps 3;  
QY 152 TCCATCATCACCACAACTCAAAGGGGTTTCCGTCACTCAGATCAAGAGCTTGGAAACCG 211  
Db 214 TCCGTACGACCAAGTCAAGGGCGTGTGACCAACTTCTTAACTTGGATTCCGG 273  
QY 212 CTGTGGGATGGCGCACTTCGTGAAGCCACTCAGGAGAGAGCTTCTTCTTGTGTG 271  
Db 274 ATCTGGGATGGCGGATTTATGTATCCAGCTCAGGAGGAAACTCCCTCTTGTGATG 333  
QY 272 ACCAACTTCTTGTGACGCGCAGCCAAAGTTCAGGCGAGATGCCAGAGACCCCTCCGTC 331  
Db 334 ACCAAGTGTATCTTCACTGATGACAGACAGAGGCTGTGCCCGAGATTC---CAGAT 390  
QY 332 CCATGGGTAACTGTGTGGGTGACGAGACTGCCCCGAAGGGAGGGAGGCACACACAGC 391

```
Db 391 GCGACACTGTGTAAATCAGATGCCAGTGTACTGCGGGTCTGCGGACCCACAGC 450
Qy 392 CACGGTGTAAACAGAGCCAGTGTGTGTTCATATGGACCCACAGGACCTGTGAGATC 451
Db 451 AACGGAGTCTCAACAGGAGTGTGTGTGTTCATATGGACCCACAGGACCTGTGAGATC 510
Qy 452 TGGAGTGTGTGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 508
Db 511 GCGGCTGTGTGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 570
Qy 509 GCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 568
Db 571 GCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 630
Qy 569 AAGTCCAAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 628
Db 631 AAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 690
Qy 629 CAATTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 688
Db 691 AAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 750
Qy 689 ACCTTCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 748
Db 751 AGTTTCCAGGACATGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 810
Qy 749 GACCTGGACACCGGGAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 806
Db 811 AACTTGGACAGAGCGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 870
Qy 807 -----AGACTACAACTTCCAGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 841
Db 871 CGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 930
Qy 842 GAGCAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 901
Db 931 GACCTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 990
Qy 902 CTCGTACCGGACGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 961
Db 991 ATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1050
Qy 962 GGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1021
Db 1051 GCGCTGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1110
Qy 1022 AGAGAAGCCCATTTCTACTGGAGGACAAAGTATGAGGAGG 1061
Db 1111 AAGAAAGACTTCTACTTCCGGAGAGAAATATAATATG 1150
```

## RESULT 6

```
US-08-742-621-2
; Sequence 2, Application US/08742621
; Patent No. 5856129
; GENERAL INFORMATION:
; APPLICANT: HILLMAN, JENNIFER L.
; APPLICANT: COLEMAN, ROGER
; TITLE OF INVENTION: NOVEL HUMAN PURINOCEPTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08742,621
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0147 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1762 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; IMMEDIATE SOURCE:
; LIBRARY: CONSENSUS
; CLONE: US-08-742-621-2
```

```
Query Match 8.9%; Score 239.2; DB 2; Length 1762;
Best Local Similarity 56.3%; Pred. No. 1.8e-51;
Matches 529; Conservative 0; Mismatches 378; Indels 33; Gaps 3;

Qy 152 TCATCATCATCAAACTCAAAAGGGTTCGTCATCATCATCAAGAGGCTTGGAAACCGG 211
Db 214 TCGGTACGACCAAGTCAAGGCGTGGTGTGACCAACTTCTAAACTTGGATTCGG 273
Qy 212 CTGTGGGATGCGCCACTTCGTGAAGCCACTCAGGAGAGAACGTGTCTTCTTGGTG 271
Db 274 ATCTGGATGTGGCGGATATGTGATACCACTCAGGAGGAAACTCCCTCTTCTGTCATG 333
Qy 272 ACCAACTTCTGTGAGCGCCAGCCCAAGTTCAAGGCGAGATGCCAGAGACACCGCTCCGTC 331
Db 334 ACCAAGTGTCTCTACCATGAACACACAGGCGCTGTGCCCGAGATTC---CAGAT 390
Qy 332 CCACTGGCTAACTGTGTGGTTCGACGAGGACTGCCCGAAGGGAGGAGGACACACAGC 391
Db 391 GCGACCACTGTGTGTAAATCAGATGCCAGTGTACTGCGGGTCTGCGGACCCACAGC 450
Qy 392 CACGGTGTAAACAGCCAGCTGTGTGTGTTCATGGGACCCACAGGACCTGTGAGATC 451
Db 451 AACGGAGTCTCAACAGGAGTGTGTGTGTTCACGCGTCCGTCAAGACGTGTGAGTG 510
Qy 452 TGGAGTGTGTGCCAGTGTGAGAGTGGC---GTTGTGCCCTCGAGGCCCTGTGCGCCAG 508
Db 511 GCGGCTGTGTGCCGCTGGAGGATGACACACAGTGCACACCTGCTTTTAAAGGCT 570
Qy 509 GCGGAGAACTTCACTGTTCATCAAAACACAGTCACTTCAGGAACTTCACTTCTCT 568
Db 571 GCAGAAACTTCACTCTTTTGGTTAAGAAACATCTGTATATCCCAAAATTAATTTTTCAGC 630
Qy 569 AAGTCCAAATGCTTGGAGACCTGGGACCCCACTATTTTAAAGCACTGCGCTATGAACA 628
Db 631 AAGGAGTATATCTTCCCAACTCACTACTTACCTCAAGTGTGTGATTTATGATGCT 690
Qy 629 CAATTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 688
Db 691 AAGGAGTATATCTTCCCAACTCACTACTTACCTCAAGTGTGTGATTTATGATGCT 750
Qy 689 ACCTTCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 748
Db 751 AGTTTCCAGGACATGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 810
Qy 749 GACCTGGACACCGGGAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 806
Db 811 AACTTGGACAGAGCGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 870
```

807 -----AGAGCTAAGCTTACAGGACAGCCACTCACTGGTG 841  
 Db CGGACGCTGAGCAACAAGTATCTCTGGCTACAAATTCAGGTTGCCAAGTACTACAGA 930  
 QY GAGCAACCGGCTGTGGAGCCCGCACCTCTCAAGCTCTATGGAATCCCGCTTCGACATC 901  
 Db GACTGCTGCAACGAGCAGCGCATCATCAAGCCCTATGCAATCCCGCTTCGACATC 990  
 QY CTGCTACCGGCGAGGAGGAAATTCGGGCTCATCCCGAGCCGCTCACTAGGAGCACC 961  
 Db ATTGTGTTGGGAAGGAGGAAATTTGACATCATCCCACTATGATCAACATCGGCTCT 1050  
 QY GGGGACGCTTGGCTGGGCGTGGTCACTTTTCTGTACCTGCTACTGCTGTATGATG 1021  
 Db GGCCTGGACCTGCTAGGACGCGTGTCTGTGTGACATCATAGTCTCTCTACTGATG 1110  
 QY 1022 AGAGAAGCCCTTTCTACTGTGAGGACAAAGPATGAGGAGG 1061  
 Db 1111 AAGAAAGACTCTACTATCGGAGAGAAATATAATATG 1150

## RESULT 7

US-09-191-608-21  
 ; Sequence 21, Application US/09191608  
 ; Patent No. 6242216  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynch, Kevin J.  
 ; APPLICANT: Burgard, Edward C.  
 ; APPLICANT: Metzger, Randy E.  
 ; APPLICANT: Niforatos, Wendie  
 ; APPLICANT: Touma, Edward B.  
 ; APPLICANT: Van Biesen, T.  
 ; TITLE OF INVENTION: Nucleic Acids Encoding a Functional  
 ; TITLE OF INVENTION: Human Purinoreceptor P2X2 and P2X4 And Methods Of Production  
 ; TITLE OF INVENTION: And Use Thereof  
 ; FILE REFERENCE: 6394.US.P1  
 ; CURRENT APPLICATION NUMBER: US/09/191.608  
 ; CURRENT FILING DATE: 1998-11-13  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 21  
 ; LENGTH: 1206  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-191-608-21

Query Match 8.8%; Score 237.6; DB 3; Length 1206;  
 Best Local Similarity 56.2%; Pred No. 3.9e-51;  
 Matches 528; Conservative 0; Mismatches 379; Indels 33; Gaps 3;  
 QY 152 TCCATCATCAACAACTCAAGGGGTTTCGGTCACTAGATCAAGGAGCTTGGAAACGG 211  
 Db 207 TCCGTTAGCAACCAAGGTCAGGGGCTGTGTGACCAACACTTCTAAACTTGGATTCCGG 266  
 QY 212 CTGCTGGATGGCCGACTTGTGAAGCCACTCAGGAGAGAGAGCTGTCTTCTTCTGGTG 271  
 Db 267 ATCTGGGATGGCGGATATATGTATACAGCTAGGAGGAAACTCCCTTCTGTCATG 326  
 QY 272 ACCAACTTCTCTGTGACGCCACCCAGTTTCAGGCGAGATGCCAGAGACCCCTCCGTC 331  
 Db 327 ACCAAGTGATCTCTCAACATGAACAGACAGAGGCGCTGTGCCCGGAGATTC---CAGAT 383  
 QY 332 CCACTGGCTAAGTCTGGTGCAGGAGACTGCCCGAGGGGAGGAGGACACACAGC 391  
 Db 384 GCGACCACTGTGTGAATACAGATCCAGCTGTACTCCGCTCTGCGGACCCACAGC 443  
 QY 392 CACGCTGTAAACAGGCGAGTGTGTGTGTTCATGAGCCACAGGACCTGTGAGATC 451  
 Db 444 AACGAGTCTCAACAGGAGGTCGTAGCTTTCACGGGTCGTCAAGACGCTGTGAGTG 503  
 QY 452 TCGAGTTGGTCCCAAGTGGAGAGTGGC---GTTGTGCCCTCGAGGCCCTCTGCCCCAG 508  
 Db 504 GCGGCTGTGTGCGGCTGGAGGATGACACACAGCTGCCAACCTGTCTTTTAAAGGCT 563

QY 509 GCCAGAACTTCAACCTGTTTCATCAAAAACACAGTCACTTTCAGCAAGTTCAACTTCTCT 568  
 Db 564 GAGAAACTTCACTCTTTTGGTTTAAAGAAACAATCTGGTATCCCAATTTAAATTTTTCAGC 623  
 QY 569 AAGTCCCAATGCTTGGAGACCTGGGACCCCACTTATTTAAGCACTGCGGCTTGAACCA 628  
 Db 624 AAGAGGAATATCTTCCCAACATCACCACCTACTTACCTCAAGTCTGTGCAATTTATGATGCT 683  
 QY 629 CAATTCAGCCCTACTGTCCTGCTTCCGCAATTTGGGACCTCTGCGGCAAGCTTGGAGGG 688  
 Db 684 AAAACAGATCCCTTCTGCCCCATATTCCTGCTTTGGCAAAATAGTGGAGAACGAGGACAC 743  
 QY 689 ACCTTCGAGGACCTGGCTTGTGGTGGCTCTGTAGGCATCAGAGTTCACTGGGATGTT 748  
 Db 744 GCTTTCAGGACATGGCGTGGAGGGAGGCATCATGGCATCCAGGTCAACTGGGACTGC 803  
 QY 749 GACCTGGACACCGGGGACTCTGGCTGGCTGCTCACTACTCTTCCAGCTCGAGGAGA-- 806  
 Db 804 AACCTGGACAGAGCCGCTCTCTGCTTGGCCAGGTACTCTTCCGCGGCTCGATACA 863  
 QY 807 -----AGAGCTAACAATTTTCAGGACAGCCACTCACTGGTG 841  
 Db 864 CGGACGCTGAGCAACAAGTATCTCTCTGCTACAAATTCAGTTTGGCAAGTACTACAGA 923  
 QY 842 GAGCAACCGGCTGTGGAGCCCGCACCTCTGCTCAAGCTCTATGGAATCCGCTTTCAGATC 901  
 Db 924 GACCTGGCTGGCAACGAGCAGCGCAGCTCATCAAGGCTATGGCATCCGCTTTCGACATC 983  
 QY 902 CTCGTACCGGCGAGGAGGAGTTCCGGCTCATCCCAAGGCTGCTCACTGGGCTGACATCGGCTCT 961  
 Db 984 ATGTGTTTGGAGGAGGAGGAAATTTGACATCATCCCACTATGATCAACATCGGCTCT 1043  
 QY 962 GGGGAGCTTGGCTGGGCTGCTCACTTTTCTGTGACCTGCTACTCTCTGTATGATG 1021  
 Db 1044 GCGCTGGCATCTCTAGGATGGGACCGCTGTGTGACATCATAGTCTCTCTACTGTCATG 1103  
 QY 1022 AGAGAAGCCCAATTTCTACTGGAGGACAAAGTATGAGGAGG 1061  
 Db 1104 AAGAAAGACTCTCTACTTATCGGAGAGAAATATAATATG 1143

## RESULT 8

US-08-750-134A-10  
 ; Sequence 10, Application US/08750134A  
 ; Patent No. 5985603  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VALERA, SOLEDAD  
 ; APPLICANT: BUELL, GARY  
 ; TITLE OF INVENTION: P2X RECEPTORS (PURINOCEPTOR FAMILY)  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: NIXON & VANDERHUYE P.C.  
 ; STREET: 1100 NORTH GLEBE ROAD  
 ; CITY: ARLINGTON  
 ; STATE: VIRGINIA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22201-4714  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/750.134A  
 ; FILING DATE: 22-JAN-1997  
 ; CLASSIFICATION: 536  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAWFORD, ARTHUR C.  
 ; REGISTRATION NUMBER: 25,327  
 ; REFERENCE/DOCKET NUMBER: 1430-116  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 816-4006

TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2643 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-750-134A-10

Query Match 8.7%; Score 233.4; DB 2; Length 2643;  
Best Local Similarity 56.7%; Pred. No. 6.9e-50;  
Matches 526; Conservative 0; Mismatches 371; Indels 30; Gaps 4;

QY 164 AAATCAAGGGGTTCCGTCTCAGATCAAGAGCTTGAAACCGGCTGTGGATGTG 223  
Db  
QY 375 AAATCAAGGGGCTGGCGTGCACCGAGCTCCCTGGCGCCCGCCAGCTGTGGATGTG 434  
Db  
QY 224 GCCGACTTCGTGAAGCCACTCAGGAGAGAGCGTTCTTCTGGTGACCACTTCCTT 283  
Db  
QY 435 GCTGACTAGCTTCCAGCCCGAGGGGACAACTCCTTGGTGATGACCAATTTCAATC 494  
Db  
QY 284 GTGACCGCAGCCCAAGTTTCAAGGCGAGATGCCCGAGAGACCGGTCCGCTCCGCTAAC 343  
Db  
QY 495 GTGACCCCGAAGCAGACTCAAGGCTACTGCGCAGAGCACCC-----AGAAGGGGGCATA 548  
Db  
QY 344 TGCTGGTTCGACGAGACTGCCCGAAGGGGAGGAGGCACACAGAGCAGCGTGTAAAA 403  
Db  
QY 549 TCGAAGGAAGCAGTGGCTGTACCCCTGGGAAGGCCAAGAGAGGCGCCAGGCGATCGC 608  
Db  
QY 404 ACAGCCGAGTGTGGTGTTCATATGGGACCCACAGAGACTGTGAGATCTGGAGTTGGTGC 463  
Db  
QY 609 ACGGCAAGTGTGGCTTCAACACACTGTGAAGAGCTGTGAGATCTTTGGCTGTGC 668  
Db  
QY 464 CAGTGGAGAGTGGCTGTGGCCCTCGAGGC---CCCTGTGGCCCGCCAGGCCAGAACTTC 520  
Db  
QY 669 CCGTGGAGGTGGATGACGACATCCCGCGCCCTGCGCTTCCGAGAGCGCCGAGAACTTC 728  
Db  
QY 521 ACACGTGTTCAACAAACACAGTCACTTCAGCAGAGTTCACTTCTTAAGTCAATGCC 580  
Db  
QY 729 ACTCTTTTCATCAAGAACAGCATAGCTTTCACGCTTCAAGGTCAAGCGGCAACCTG 788  
Db  
QY 581 TTGAGAGCTGGGACCCCACTTATTTAAGCACTCCGCTATGAACCAATTCAGCCCC 640  
Db  
QY 789 GTGAGGAGGTGAATGCTGCCACATGAAGACTTCCCTTTTCAAGACCTTCGACCCC 848  
Db  
QY 641 TACTTCCCGTTCGGCATTTGGGACCTCTGGCCAGGCTGGAGGACCTTCGAGGAC 700  
Db  
QY 849 CTGTGCCAGCTTCCAGCTTGGCTACGTTGGTGCAGAGTCAAGGCTCAGGCACTTCAGCAC 908  
Db  
QY 701 CTGGCGTTGCTGGGTGGCTCTGTAGGCATCAGAGTTCACTGGGATTTGACCTGGACACC 760  
Db  
QY 909 CTGGCTGAGAGGGTGGAGTGGTTGGCATCACCATCGACTGGCACTGTGACTGGACTGG 968  
Db  
QY 761 GGGGACTCTGGTCTGGCGCTCACTACTCTTTCAGCTGAGGAGAGA----- 809  
Db  
QY 969 CAGGTACGCACTGAGAGCCCACTATGAGTTCCATGGGCTGTACGAGAGAAAAATCTC 1028  
Db  
QY 810 -----GCTACAACTTCAGAGACCCACTCACTGGTGGGAGCAACCGGTGTGGAGGCC 862  
Db  
QY 1029 TCCCGAGGCTTCACTTCAAGTTTCCAGGACACTTTTGGAGAAC---GGACCAACTAC 1085  
Db  
QY 863 CCACCCCTGCTCAAGCTATGGAATCCGCTTCGACATCTCTGTCACCGGCGAGCGAGGG 922  
Db  
QY 1086 CGTCACTCTTCAAGGTGTTGGGATTCGCTTTCATCTCTGCTGGAGCGGCAAGGCCGGG 1145  
Db  
QY 923 AGTTTGGGCTCATCCCGAGCCGCTGCACACTGGGACCGGGGAGCTTGGCTGGGGTG 982  
Db  
QY 1146 AGTTTGCATCATCCCTCAATGACCACTTCGCTCTGGAATTTGGATCTTTGGGGTG 1205  
Db  
QY 983 GTCACTTTTCTGTGACCTGCTACTGCTGTATGTGATAGAGAGGCCAATTTCTACTGG 1042  
Db  
QY 1206 GCCACAGTTCTGTGACCTGCTGCTTTCATCTCTGCTTCACTCTGCTTAAGAGGCACTACTACAAG 1265

QY 1043 AGGACAAAGTATGAGGAGGCGCAAGGCC 1069  
Db 1266 CAGAAGAAGTTCAATACGCTGAGGAC 1292

RESULT 9  
US-09-363-745-10  
; Sequence 10, Application US/09363745  
; Patent No. 6194162  
; GENERAL INFORMATION:  
; APPLICANT: VALERA, SOLEDAD  
; APPLICANT: BUELL, GARY  
; TITLE OF INVENTION: P2X RECEPTORS (PURINOCCEPTOR FAMILY)  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHUYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/363,745  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/750,134  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAWFORD, ARTHUR C.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 1430-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4006  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2643 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-363-745-10

Query Match 8.7%; Score 233.4; DB 3; Length 2643;  
Best Local Similarity 56.7%; Pred. No. 6.9e-50;  
Matches 526; Conservative 0; Mismatches 371; Indels 30; Gaps 4;

QY 164 AAATCAAGGGGTTCCGTCTCAGATCAAGAGCTTGAAACCGGCTGTGGATGTG 223  
Db 375 AAATCAAGGGGCTGGCGTGCACCGAGCTCCCTGGCGCCCGCCAGGCTGTGGATGTG 434  
QY 224 GCCGACTTCGTGAAGCCACTCAGGAGAGAGCGTTCTTCTTGGTGACCACTTCCTT 283  
Db 435 GCTGACTAGCTTCCAGCCCGAGGGGACAACTCCTTGGTGATGACCAATTTCAATC 494  
QY 284 GTGACCGCAGCCCAAGTTTCAAGGCGAGATGCCCGAGAGACCGGTCCGCTCCGCTAAC 343  
Db 495 GTGACCCCGAAGCAGACTCAAGGCTACTGCGCAGAGCACCC-----AGAAGGGGGCATA 548  
QY 344 TGCTGGTTCGACGAGACTGCCCGAAGGGGAGGAGGCACACAGGCGAGCGTGTAAAA 403  
Db 549 TCGAAGGAAGCAGTGGCTGTACCCCTGGGAAGGCCAAGAGAGGCGCCAGGCGATCCGC 608  
QY 404 ACAGCCGAGTGTGGTGTTCATATGGGACCCACAGAGACTGTGAGATCTGGAGTTGGTGC 463  
Db 609 ACGGCAAGTGTGGCTTTCATCTCTGCTTCAAGAGCGTGTGAGATCTTTGGCTGTGC 568

QY 464 CCAGTGGAGAGTGGCGTTGTGCTCCGAGG-CCCTGCTGGCCAGGCGCCAGAACTTC 520  
DB |||||  
QY 669 CCCGTGGAGTGGATGACGACATCCCGCGCCCTGCTCCAGAGGCGGAACTTC 728  
DB |||||  
QY 521 ACATGTTTCATCAAAAACACAGTACCTTCAGCAAGTTCATCTCTTAAGTCCAAATGCC 580  
DB |||||  
QY 729 ACTCTTTTCATCAAGAACAGATCAGCTTTCCAGGCTTCAGGTCAACAGGCGCAACTG 788  
DB |||||  
QY 581 TTGAGACCTGGAGCCCGACCTATTTTAAAGCACTGCGCTATGAACACAAATTCAGCCCC 640  
DB |||||  
QY 789 GTGAGGAGGTGAATGCTGCCCAATGAGACCTGCTCTTCAAGACCTTCAGCCCC 848  
DB |||||  
QY 641 TACTGTCCCGTTCGCGATTTGGGAGCTCTGTGCGCAAGGCTGGAGGACCTTCGAGGAC 700  
DB |||||  
QY 849 CTGTGCCAGTCTTCCAGCTTGGCTACGTGTGCAAGAGTCAGGCCAGAACTTCAGCACC 908  
DB |||||  
QY 701 CTGGCTTCTGGGTGGCTCTGTAGGATCAGAGTTCATCTGGGATTTGACCTCGACACC 760  
DB |||||  
QY 909 CTGGCTGAGAGGTGGAGTGGTGGCATCACCATCGACTGGCCTGTGACCTGGACTGG 968  
DB |||||  
QY 761 GGGGACTCTGGCTGGCTCTACTCTCTTCCAGCTGCGAGGAGAGA----- 809  
DB |||||  
QY 969 CACGTACGGCACTGACAGACCCATCTATGAGTTCCATGGCTGTACGAAGAGAAAATCTC 1028  
DB |||||  
QY 810 -----GCTACAACTTCAGACAGCCACTCACTGTTGGGAGCAACCGGTGTGGAGGCC 862  
DB |||||  
QY 1029 TCCCAGGCTTCAACTTCAGGTTTGGCAGGCACCTTTTGTGGAGAAC---GGGACCACTAC 1085  
DB |||||  
QY 863 CGCACCTTCTCAAGCTCTATGGAATCCGCTTCGACATCTCTGTCACCGGCGAGGAGGG 922  
DB |||||  
QY 1086 CGTCACTCTTCAAGTGTGGGTTTGGCTTTCATCTCTGTTGGAGCGCAAGGCGGG 1145  
DB |||||  
QY 923 AAGTTCGGGCTCATCCCAAGCGCGTCACTGCGGACCGGCGAGCTTGGCTGGGCGTG 982  
DB |||||  
QY 1146 AAGTTTGACATCATCCCTACAAATGACCACCATCGGCTCTGGAATTTGGAGTCT 1205  
DB |||||  
QY 983 GTCACTTTTCTGTGACTCTACTGCTGTATGTGATAGAGAGCCCAATTTCTACTGG 1042  
DB |||||  
QY 1206 GCGCAGTCTCTGTGACTCTGCTGCTTCAATCTGCTTACGCTAGAGGCACTACTACAG 1265  
DB |||||  
QY 1043 AGGACAAAGTATGAGGCGCAAGGCC 1069  
DB |||||  
QY 1266 CAGAAGAGTTCAAATACGCTGAGGAC 1292  
DB |||||

## RESULT 10

US-09-023-655-897  
Sequence 897, Application US/09023655  
Patent No. 6607879  
GENERAL INFORMATION:  
APPLICANT: Cocks, Benjamin G.  
APPLICANT: Susan G. Stuart  
APPLICANT: Jeffrey J. Sellhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: word perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,655  
FILING DATE: HERewith  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 897:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2643 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g1166437  
US-09-023-655-897  
Query Match 8.7%; Score 233.4; DB 4; Length 2643;  
Best Local Similarity 56.7%; Pred. No. 6.9e-50;  
Matches 526; Conservative 0; Mismatches 371; Indels 30; Gaps 4;  
QY 164 AAATCAAGGGGTTTCGTCATCAGATCAAGAGCTTGGAAACCGGCTGTGGATGTG 223  
DB |||||  
QY 375 AAATCAAGGGGCTGGCGGTGACCCAGCTCCCTCGCCCGCCAGGTCTGGATGTG 434  
DB |||||  
QY 224 GCGCACTTCGTGAAGCCACCTCAGGAGAGAAAGTTCCTTCTTGGTGACCACTTCCTT 283  
DB |||||  
QY 435 GCTGACTAGTCTTCCAGCCAGGGGAGCAACTCTTCTGTTGTCATGACCAATTTGATC 494  
DB |||||  
QY 284 GTGACCGCAGGCCAAGTTTCAGGGCAGATGCCCAGAGACCCGTCGTCGCCACTGGTAAAC 343  
DB |||||  
QY 495 GTGACCCCGAAGCAGACTCAAGGCTACTGCGCAGAGCACCC-----AGAAGGGGGGATA 548  
DB |||||  
QY 344 TGCTGGGTGCGAGGAGTCCCGGAGGGGAGGAGGAGCACACACCGCCGTTGTAATAA 403  
DB |||||  
QY 549 TCAAGGAGAGACAGTGGCTGTACCCCTGGGAGGCGCAAGAGAGGCGCCAAAGGATCCGC 608  
DB |||||  
QY 404 ACAGGCCAGTGTGTGCTTCAATGGGACCCACAGGACCTGTGAGATCTGAGTTGGTGC 463  
DB |||||  
QY 609 ACGGCAAGTGTGTGCTTCAACGACACTGTGAAGAGCTGTGAGATCTTGGCTGGTGC 668  
DB |||||  
QY 464 CCAGTGGAGAGTGGCGTTGTGCTCCGAGG-CCCTGCTGGCCAGGCGCCAGAACTTC 520  
DB |||||  
QY 669 CCCGTGGAGTGGATGACGACATCCCGCGCCCTGCTCCAGAGGCGGAACTTC 728  
DB |||||  
QY 521 ACATGTTTCATCAAAAACACAGTCACTTCAGCAAGTTCATCTCTTAAGTCCAAATGCC 580  
DB |||||  
QY 729 ACTCTTTTCATCAAGAACAGATCAGCTTCAGCTTCAGGTCAACAGGCGCAACTG 788  
DB |||||  
QY 581 TTGAGACCTGGAGCCCGACCTATTTTAAAGCACTGCGCTATGAACACAAATTCAGCCCC 640  
DB |||||  
QY 789 GTGAGGAGGTGAATGCTGCGCCACATGAAGACCTGCTCTTTCACAGACCCCTGCACCCC 848  
DB |||||  
QY 641 TACTGTCCCGTTCGCGATTTGGGAGCTCTGTGCGCAAGGCTGGAGGACCTTCGAGGAC 700  
DB |||||  
QY 849 CTGTGCCAGTCTTCCAGCTTGGCTACGTGTGCAAGAGTCAGGCCAGAACTTCAGCACC 908  
DB |||||  
QY 701 CTGGCTTCTGGGTGGCTCTGTAGGATCAGAGTTCATCTGGGATTTGACCTCGACACC 760  
DB |||||  
QY 909 CTGGCTGAGAGGTGGAGTGGTGGCATCACCATCGACTGGCCTGTGACCTGGACTGG 968  
DB |||||  
QY 761 GGGGACTCTGGCTGGCTCTACTCTCTTCCAGCTGCGAGGAGAGA----- 809  
DB |||||  
QY 969 CACGTACGGCACTGACAGACCCATCTATGAGTTCCATGGCTGTACGAAGAGAAAATCTC 1028  
DB |||||  
QY 810 -----GCTACAACTTCAGACAGCCACTCACTGTTGGGAGCAACCGGTGTGGAGGCC 862  
DB |||||  
QY 1029 TCCCAGGCTTCAACTTCAGGTTTGGCAGGCACCTTTTGTGGAGAAC---GGGACCACTAC 1085  
DB |||||

QY 863 GCACCCCTGCTCAAGCTCTATGGAATCGCTTCGACATCCTCGTCAACCGGCGCAGCAGG 922  
Db 1086 CGTACCTCTTTCAAGGTGTTGGGATTCGCTTTGACATCCTGGTGGACGCAAGGCCGGG 1145  
QY 923 AAGTTGGGCTCATCCCAAGCGCTGACACTGGGCGACGGGCGAGCTTGGCTGGGGTG 982  
Db 1146 AAGTTTGACATCATCTTACATGACACCATCGGCTCTGGAATTGGCATCTTTGGGGTG 1205  
QY 983 GTACCTTTTCTGTGACCTGCTACTGCTGTATGTGATGATGAGAGGCCATTTCTACTGG 1042  
Db 1206 GCCACAGTTCTGTGACCTGCTGCTGTCTGCTGTCTGACATCTCTGCTTAAGAGGCACTACTACAAG 1265  
QY 1043 AGGACAAAGTATGAGAGGCCAAGGCC 1069  
Db 1266 CAGAAGAAGTTCAATACGCTGAGGAC 1292

RESULT 11  
US-08-750-134A-5  
; Sequence 6, Application US/08750134A  
; Patent No. 5985603  
; GENERAL INFORMATION:  
; APPLICANT: VALERA, SOLEDAD  
; APPLICANT: BUELL, GARY  
; TITLE OF INVENTION: P2x RECEPTORS (PURINOCEPTOR FAMILY)  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/750,134A  
; FILING DATE: 22-JAN-1997

CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAWFORD, ARTHUR C.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 1430-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4006  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1997 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdNA  
US-08-750-134A-6

Query Match 8.4%; Score 225.4; DB 2; Length 1997;  
Best Local Similarity 55.1%; Pred. No. 6.8e-48;  
Matches 548; Conservative 0; Mismatches 411; Indels 36; Gaps 4;  
QY 97 GTGGGCTCTCTGCGCAAAAGGTTACCGAGGCGGGACCTGGAAACCCAGTTTCCAT 156  
Db 235 GTGGGCTCTCTGCGCAAAAGGTTACCGAGGAAACCGGACTCCGTGGTC--AGCTCGGT 291  
QY 157 CATCAACAACTCAAGGGGTTCCGCTACTCAGATCAAGGAGCTTGGAAACCGGCTGTG 216  
Db 292 GACAAACAAAGCAAGGTTGGCTGTGACCAACCTCTCAGCTTGGATTCGGATCTG 351  
QY 217 GGATGTGGCGGACTTCGTGAAGCCACTCAGGAGAGAACGTGTTCTTCTTGGTACCAA 276

RESULT 12  
US-09-363-745-6  
; Sequence 6, Application US/09363745  
; Patent No. 6194162  
; GENERAL INFORMATION:  
; APPLICANT: VALERA, SOLEDAD  
; APPLICANT: BUELL, GARY  
; TITLE OF INVENTION: P2x RECEPTORS (PURINOCEPTOR FAMILY)  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA

Db 352 GGACGTGGCGGACTATGTGATTCACGCTCAGGAGGAAAACTCCCTCTTCATTATGACCAA 411  
QY 277 CTTCTTTGTGACCCAGCCCAAGTTTCAGGGCAGATGCCAGAGCACCCGCTCCGCTCCCACT 336  
Db 412 CATGATGTGCACCGTGAACCCAGACACAGAGACCTGTCCAGAGATTC---CTGATAGAC 468  
QY 337 GGCTAACTGTGGGTGACAGGACTGCCCGAAGGGAGGAGGAGGACACACAGCCACCG 396  
Db 469 CAGCATTTGTAATTCAGACGCCGACTGCACTCTCTGGCTCCGTGGACACCCAGCAGTGG 528  
QY 397 TGTAAAAACAGGCGAGTGTGTGTTCAATGGGACCCACAGGACCTGTGAGATCTGGAG 456  
Db 529 AGTTGGCACTGGAAGATGTGTTCTTCAATGAGTCTGTGAAGACCTGTGAGGTGGCTGC 588  
QY 457 TTGGTGCCAGTGGAGAGTGGCGTT---GTGCCCTCGAGGCCCTCTCTGGCCCGAGCCCA 513  
Db 589 ATGCTGCCCGGTGGAGAACGACGTTGGGTGGTGGCCACGCCGCTTCTTAAAGGCTCGAG 648  
QY 514 GAACTTTCACACTGTTTCATCAAAAACACAGTCACCTTCAGCAAGTTCAACTTCTTAAGTC 573  
Db 649 AAACTTCACCTCTTGTGTAAGAAACACATCTGGTACCCCAAGTTTAACTTCAGCAAGAG 708  
QY 574 CAATGCTTGGAGACCTGGGACCCCACTTATTTTAAGCACTCGCGCTATGAAACCAATT 633  
Db 709 GAACATCTCTCCCAACATCACACGCTCTACCTCAAAATCGTGCAATTTACATGCTCAAC 768  
QY 634 CAGCCCTACTGTCCCGTGTTCGCGATTGGGGACCTCGTGGCCAAAGGCTGGAGGACCTT 693  
Db 769 GGATCCCTCTGCCCCATATTCGCTCTTGSCACAACTCGTGGGGACGCGGACATAGCTT 828  
QY 694 CGAGGACCTGGCGTGTGGTGGCTCTGTAGGACATCAGAGTTCACTGGGATTTGACCT 753  
Db 829 CCAGGAGATGGCAGTTGAGGGAGGCATCATGGGTATCCAGATCAAGTGGGACCTGCAACT 888  
QY 754 GGACACCGGGACTCTGGCTGTGGCTCACTACTCTCTCCAGCTGCAGGAGAAGA---809  
Db 889 GGATAGAGCCGCTCCCTTTCCTCGGGGCTGCACACCCGGGA 948  
QY 810 -----GTTCAACTTCAGSACAGCCACTCACTGTGGGAGCA 846  
Db 949 CTTGGAACACATGTGTCTCTCGCTACAAATTTTCAGTTTGCCTTTCAGGAGCACT 1008  
QY 847 ACCGGTGTGGAGCCCGCACCTGCTCAGCTCTATGGAATCCGCTTCGACATCTCTCGT 906  
Db 1009 GGGCGGCAAGAGAGCGGCACACTCACCAGGCGTACGGCATCCGCTTTGACATCATCGT 1068  
QY 907 CACCGGCGAGGCGAGGAAAGTTCCGGGCTCATCCCGCGCGCTCACACTGGGCGCCGGGC 966  
Db 1069 GTTTGGAAGGCTGGGAAGTTTGACATCATCCCTACCATGATCAACGTTGGCTCTGGCTT 1128  
QY 967 AGCTTGGCTGGCGTGTGTACCTTTTCTGTGACCTGCTACTGCTGTATGTGATAGAGA 1026  
Db 1129 GGGCTCTCTCGGGTGGCGACGGTGTCTCTGACGTCATAGTCTCTTACTGATGAAGAA 1188  
QY 1027 AGCCCATTTCTACTGGAGGACAAAGTATGAGGAGG 1061  
Db 1189 GAAATACTACTACCGGACAAAGATATAAGTATG 1223



COUNTRY: U.S.A.  
 ZIP: 22201-4714  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/363,745  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/750,134  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CRAWFORD, ARTHUR C.  
 REGISTRATION NUMBER: 25,327  
 REFERENCE/DOCKET NUMBER: 1430-116  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 816-4006  
 TELEFAX: (703) 816-4100  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1997 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-09-363-745-6

Query Match 8.4%; Score 225.4; DB 3; Length 1997;

Best Local Similarity 55.1%; Pred. No. 6.8e-48; Indels 36; Gaps 4;  
Matches 548; Conservative 0; Mismatches 411;

QY 97 GTGGCTCTCTCGCCAAAGGCTACAGAGGGGACCTGGAACCCAGATTTCAT 156  
 DB 235 GTGGGTGTCGTGGGAAAGGGCTACAGAAACGGACTCCGTGTC---AGTCGGT 291  
 QY 157 CATCAACAACTCAAGGGTTTCGTCACTCAGATCAAGAGCTTGAAACCGGCTGTG 216  
 DB 292 GACAAACAAAGGCTGGGTGTGACCAACACCTCTCAGCTTGGAATCCGGATCTG 351  
 QY 217 GGATGTGGCCGACTTCGTGAAGCCACCTCAGGGAGAGAGCTGTCTTCTGTGACCA 276  
 DB 352 GGACGTGGCGACTATGATGATTCAGCTCAGAGGAACTCCCTCTTCAATTATGACCA 411  
 QY 277 CTTCTTGTGACGCCAGCCCAAGTTACGGGAGATGCCCAGAGCACCCGCTCCCACT 336  
 DB 412 CATGATTGTCACTGACCAACAGACAGAGACCTGTCCAGAGATTC---CTGATAGAC 468  
 QY 337 GGCTAACTGCTGGTGCAGAGGACTGCCCGAAGGGAGGGAGGACACACAGCCACGG 396  
 DB 469 CAGCATTTGTAATTCAGACGCGGACTGCATCTCTGGCTCCGTGGACACCCACAGCAGTGG 528  
 QY 397 TGTAATAACAGCCAGTGTGTGTGTTCAATGGGACCCACAGGACCTGTGAGATCTGGAG 456  
 DB 529 AGTTGCGACTGGAAGATGTGTCTTCAATGAGTCTGTGAGACCTGTGAGGTGGCTGC 588  
 QY 457 TTGTGCGCCAGTGAGAGTGGCGTT---GTGCGCTCAGAGCCCTGTGCGCCAGGCCCA 513  
 DB 589 ATGGTGGCCGTGGAGAACGACGTGTGGCGTGCCAAACCGCGCTTCTTAAAGGCTGAGA 648  
 QY 514 GAACTTCACACTGTTCATAAAACACAGTCACTTCAGCAAGTTCACTTCTCTAAGTC 573  
 DB 649 AACTTTCACCTCTTGTGAAGACACATCTGGTACCCCAAGTTTAACTTCAGCAGAG 708  
 QY 574 CAATGCTTGGAGACTTGGGACCCCACTATTATTAAGCACTGCGGTATGAACCAAAAT 633  
 DB 709 GAACATCTCTCCCAACATCACCAGCTCTACCTCAATCTGTGATTTACAATGCTCAAA 768  
 QY 634 CAGCCCTACTGTCCCGTGTTCGGATTTGGGACCTGTGCGCCCAAGGCTGGGGACCTT 693  
 DB 769 GGATCCCTCTTGCCCAATATTCGCTTTGGCAAACTCGTGGGAGACCGGGACATAGCTT 828

QY 694 CGAGGACCTGGCGTTCCTGGTGGCTCTCTAGGCATCAGAGTTCACTGGGATTGTGACCT 753  
 DB 829 CCAGGAGATGGCAGTTTGAAGGGAGGCATCATGGGTATCCAGATCAAGTGGGACTGCAACT 888  
 QY 754 GGACACCGGGGACTCTGGGTGTCTGGGCTCACTACTCTTCCAGCTGCAGGAGAAGA--- 809  
 DB 889 GGATAGAGCGGCTCCCTTGGCTGCCAGATATTCCTTCGGCGCTGACACCCGGGA 948  
 QY 810 -----GCTACAACTTCAGGACGCACTCACTGGTGGGAGCA 846  
 DB 949 CTTGGAACACAATGTGTCTCTGGGTACAAATTCAGGTTTGGCAAGTACTACAGGAGCT 1008  
 QY 847 ACGGGTGTGGAGCGCCGACCTCTCAAGCTCTATGGAATCCGTTTGGATCCTCGT 906  
 DB 1009 GGCCGCAAGAGCAGCGCACACTCACAAAGGCTACGSCATCCGCTTGGATCATCGT 1068  
 QY 907 CACCGGCGAGGAGGAAGTTCCGGCTCATCCCAAGCGCGTCACTGGGACCGGGGC 966  
 DB 1069 GTTTGAAAAGGCTGGGAAGTTTGACATCATCCCTACCATGATCAACGTTGGCTCTGGCTT 1128  
 QY 967 AGCTTGGCTGGCGGTGTCACTCTTTCTGTGACCTGTCTGCTATGTGATGATAGAGA 1026  
 DB 1129 GCGCTCTCGGGGTGGCGACGGTCTCTGTGACGTCTATGCTCTACTGATGATGAGAA 1188  
 QY 1027 AGCCCATTTCTACTGGAGGACAAAGTATGAGGAGG 1061  
 DB 1189 GAAATACTACTACCGGGACAGAAATATAAGTATG 1223

## RESULT 13

US-09-191-608-14  
 ; Sequence 14, Application US/09191608  
 ; Patent No. 6242216  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynch, Kevin J.  
 ; APPLICANT: Burgard, Edward C.  
 ; APPLICANT: Metzger, Randy E.  
 ; APPLICANT: Niforatos, Wende  
 ; APPLICANT: Touma, Edward B.  
 ; APPLICANT: Van Biesen, T.  
 ; TITLE OF INVENTION: Nucleic Acids Encoding a Functional  
 ; TITLE OF INVENTION: Human Purinoreceptor P2X2 and P2X4 And Methods Of Production  
 ; FILE OF INVENTION: And Use thereof  
 ; FILE REFERENCE: 6394.US.PI  
 ; CURRENT APPLICATION NUMBER: US/09/191,608  
 ; CURRENT FILING DATE: 1998-11-13  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 14  
 ; LENGTH: 1421  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-191-608-14

Query Match 8.0%; Score 214.4; DB 3; Length 1421;  
 Best Local Similarity 55.0%; Pred. No. 3.8e-45;  
 Matches 553; Conservative 1; Mismatches 407; Indels 45; Gaps 5;

QY 92 GAGAAAGTGGGCTCTCCTCGCCAAAAAGGCTTACAGGAGCGGGACCTGGAAACCCAGTTT 151  
 DB 174 GTGTGGTACGTATTTCATCGTCARAAAAGCTTACAGGAGAGCGAGCGGGCCCGAGAGC 233  
 QY 152 TCCATCATCAACAACTCAAAGGGTTTCCCTCACTAGATCAGAGGAGCTTGGAAACCGG 211  
 DB 234 TCCATCATCAACAAAGTCAAGGGATCACCAAGTCCACAGTCCGAGACAAAG----- 279  
 QY 212 CTGTGGGATGTGGCGGACTTCGTGAAGCCACTTCAGGGAGAGAACTGTTCTTCTTGGTG 271  
 DB 280 -TGTGGGAGTGGAGGAGTACGTGAAGCCCCCGAGGGGGGCGAGCGTGTTCAGCATCATC 338  
 QY 272 ACCAACTCTTGTGACCGCCCAAGTTTCAGGGGAGATGCCAGAGACCCCGTCGTC 331



339 ACCAGGGTCGAGGCCACCACTCCAGAGCCAGGGAACCTGCCCGAGAGCATAAGGGTC 398  
QY 332 CCACTGGCTACTGCTGGTGCAGAGGACTGCCCGAGGAGGGAGGACACACAGC 391  
Db 339 CACAAGCCGACCTGCTTCCGAGCCGACTGCGTGGGAGCTGGACATGCTGGGA 458  
QY 392 CACGGGTGTAACAAACAGGCGAGTGTGTG---GTGTTCAATGGGACCCACAGGACCTGTGAG 448  
Db 459 AACGGACCTGAGGACCGGGCGTGTGTGCTTATACAGAGGCGCTCCAGACCTGGAG 518  
QY 449 ATCTGGAGTGTGTCGCCAGTGGAGAGTGGGTGTGTGCTGAGGCGCTCTGTCGGTACGATG 508  
Db 519 GTGTTGGTGTGTCGCCAGTGGAGATGGGCTCTGTGACCAATTTCTGGGTACGATG 578  
QY 509 GCCCAGAACTTACACTGTTCAATCAAAACACAGTCACTTCCAGCAAGTTCAACTTCTCT 568  
Db 579 GCCCAGAACTTACACTGTTCAATCAAAACACAGTCACTTCCAGCAAGTTCAACTTCTCT 638  
QY 569 AAGTCCAACTGCTTGGAGACCTGGAGCCCACTTATTTAAGCACTGCGCTATGAACCA 628  
Db 639 AAGGGCAACATCGCCGACCGCAGACA---CGGGTACCTGAAGCGCTGCACGTTCCACGAG 695  
QY 629 CAATTCAGCCCTACTGTCCTGTGTCGATTTGGGACCTCGTGGCCAGGCTGGAGGG 688  
Db 696 GCTTCGACCTTACTGTCCTTCAAGCTGGGCTTTTATCGTGGAGAGCTGGGAG 755  
QY 689 ACCTTCGAGGACCTGCGCTTGTGCTGGGCTCTGTAGGCATCAGAGTTCACCTGGGATGT 748  
Db 756 AGCTTCAGAGCTCGCACAGAGGTTGTGTGTCATCGGGTCAATTCAGCTGGAGTGT 815  
QY 749 GACCTGGACACCGGGGACTCTGGCTGCTGAGCTCTACTCTTCCAGCTGCGAGGAAG 808  
Db 816 GACCTGGAGCTCGCACAGAGGTTGTGTGTCATCGGGTCAATTCAGCTGGAGTGT 875  
QY 809 A-----GCTCAACTTCAGGACAGCCACTCACTGTGTGGGAGCAA 847  
Db 876 AAGCAGCTGCTGCTCGTCAAGCTTCACTTTCAGCTTGGCA---AATCTACAAGATC 932  
QY 848 CCGGTGTGAGGCGCCGACCTGCTGCTCAAGCTCTATGAAATCCGCTTCCAGATCTGTC 907  
Db 933 AATGGCACCACACCGGACGCTCATCAAGGCTTACGGGATCCGCAATTCAGCTCAATGTG 992  
QY 908 ACCGGCAGGAGGAGTTCGGGCTCATCCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 967  
Db 993 CATGGACAGGCGGGAGTTACGCTGATTCACACATTTAATCTGGCCACAGCTGTG 1052  
QY 1028 GCCCATTTCTAGGAGGACAAAGTATGAGGAGGCGCAAGGCGCCGA 1073  
Db 1113 AACAGGCTCTACAGCCATAAGAAATTTGACAAGGTGTGTACGCCGA 1158

RESULT 14  
US-09-191-608-13  
; Sequence 13, Application US/09191608  
; Patent No. 6242216  
; GENERAL INFORMATION:  
; APPLICANT: Lynch, Kevin J.  
; APPLICANT: Burgard, Edward C.  
; APPLICANT: Metzger, Randy E.  
; APPLICANT: Niforatos, Wende  
; APPLICANT: Touma, Edward B.  
; APPLICANT: Van Biesen, T.  
; TITLE OF INVENTION: Nucleic Acids Encoding a Functional  
; TITLE OF INVENTION: Human Purinoreceptor P2X2 and P2X4 And Methods Of Production  
; TITLE OF INVENTION: And Use Thereof  
; FILE REFERENCE: 6394 US P1  
; CURRENT APPLICATION NUMBER: US/09/191,608  
; CURRENT FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13  
; LENGTH: 1436  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-191-608-13  
  
Query Match 7.9%; Score 213.2; DB 3; Length 1436;  
Best Local Similarity 55.0%; Pred. No. 7.8e-45;  
Matches 553; Conservative 0; Mismatches 408; Indels 45; Gaps 5;  
  
QY 92 GAGAAGTGGCTCTCTCCGCAAAAGGCTTACCAGAGCGGACCTGGAAACCCAGTTT 151  
Db 174 GTGTGTGATGATTCATGCTGAGAAAGCTTACAGAGAGCGAGACGGGCCCGAGAGC 233  
QY 152 TCCATCATCACCAAACTCAAAAGGGTTCGCTCACTCAGATCAAGAGCTTGGAAACCG 211  
Db 234 TCCATCATCACCAAGTCAAGGGATCACCACTCCGAGCACAAAG----- 279  
QY 212 CTGTGGATGTGCGGACTTTCGTGAAGCCACTTCAGGAGAGAACGTGTCTCTTGGTG 271  
Db 280 -TGTGGGACGTGGAGGATACGTGAAGCCCCCGAGGGGGGAGCGTGTTCAGCATCATC 338  
QY 272 ACCAACTTCTGTGAGCGCCAGCCCAAGTTTCAGGCGAGATGCCAGAGCACCCGCTCGTC 331  
Db 339 ACCAGGTTCAGGCGCACCACTCCAGACCCAGAGGAACCTGCCCGAGAGCATAGGGTC 398  
QY 332 CCACTGGCTTAATCTGGGTTCAGAGGACTGCCCCGAAGGGAGGAGGACACACAGC 391  
Db 399 CACAACCCACCTGCTCTCCGACGCCGACTGCGTGGCTGGGAGCTGGACATGCTGGGA 458  
QY 392 CAGGTGTAAACACAGCCAGTGTGTG---GTGTTCAATGGGACCCACAGAGCTGTGAG 448  
Db 459 AACGCTGAGGACCGGGCGCTGTGTGCTTATTCAGAGGGCCCTCCAGAGACTGGAG 518  
QY 449 ATCTGGAGTGTGTCGCCAGTGGAGAGTGGCTGTGTGCTGAGGCGCTGCTGGGCCAG 508  
Db 519 GTGTTGGCTGTGCTGCGGTGGAGATGGGCGCTCTGTACAGCAATTTCTGGGTACGATG 578  
QY 509 GCCAGAACTTCACTGTTTCATCAAAACACAGTCACTTTCAGAGAGTTCACTTCTCT 568  
Db 579 GCCCAAAATTTCACTTCTCATCAAGACAGATCACCTTACCCCAAAATTCCTCTTCC 638  
QY 569 AAGTCCAACTGCTTGGAGACTGGGACCCCACTTATTTAAGCACTGCCCTATGAACCA 628  
Db 639 AAGGGCAACATCGCCGACCGCAGACA---CGGGTACCTGAAGCGCTGCACGTTCCACGAG 695  
QY 629 CAATTCAGCCCTACTGTCCTGTGTCGATTTGGGACCTCGTGGCCAGAGCTGGAGGG 688  
Db 696 GCTTCGACCTTACTGTCCTTCAAGCTGGGCTTTTATCGTGGAGAGCTGGGAG 755  
QY 689 ACCTTCGAGGACCTGCGCTTGTGCTGGGCTCTGTAGGCATCAGAGTTCACCTGGGATGT 748  
Db 756 AGCTTCAGAGCTCGCACAGAGGTTGTGTGTCATCGGGTCAATTCAGCTGGAGTGT 815  
QY 749 GACCTGGACACCGGGGACTCTGGCTGCTGAGCTTCACTTCTTCCAGCTGCGAGGAAG 808  
Db 816 GACCTGGAGCTGCTGATCGAGTGCACCCCAAGTACTCTTCCGAGGCTTGAACCC 875  
QY 809 A-----GCTCAACTTCAGGACAGCCACTCACTGTGTGGGAGCAA 847  
Db 876 AAGCAGCTGCTGCTCGTCAAGCTTCACTTTCAGCTTGGCA---AATCTACAAGATC 932  
QY 848 CCGGTGTGAGGCGCCGACCTGCTGCTCAAGCTCTATGAAATCCGCTTCCAGATCTGTC 907  
Db 933 AATGGCACCACACCGGACGCTCATCAAGGCTTACGGGATCCGCAATTCAGCTCAATGTG 992  
QY 908 ACCGGCAGGAGGAGTTCGGGCTCATCCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 967  
Db 993 CATGGACAGGCGGGAGTTACGCTGATTCACACATTTAATCTGGCCACAGCTGTG 1052  
QY 1028 GCCCATTTCTAGGAGGACAAAGTATGAGGAGGCGCAAGGCGCCGA 1073  
Db 1113 AACAGGCTCTACAGCCATAAGAAATTTGACAAGGTGTGTACGCCGA 1158

Db 1053 ACTTCGTCGGGGTGGGCTCCTTCGTGCGACTGATCTTCTAATCATGACAAA 1112  
QY 1028 GCCCATTTCTACTGGAGGCAAAAGTATGAGGAGGCCAAGGCCCGA 1073  
Db 1113 AACAGGTCTACAGCCATAAGAAATTTGACAAGGTGTGTACGCCGA 1158

RESULT 15  
US-08-750-134A-4  
; Sequence 4, Application US/08750134A  
; Patent No. 5985603  
; GENERAL INFORMATION:  
; APPLICANT: VALERA, SOLEDAD  
; APPLICANT: BUELL, GARY  
; TITLE OF INVENTION: P2x RECEPTORS (PURINOCEPTOR FAMILY)  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/750,134A  
; FILING DATE: 22-JAN-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAWFORD, ARTHUR C.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 1430-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4006  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1837 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: rat P2x from vas deferens  
US-08-750-134A-4

Query Match 7.6%; Score 206; DB 2; Length 1837;  
Best Local Similarity 54.1%; Pred. No. 6.3e-43;  
Matches 500; Conservative 0; Mismatches 400; Indels 24; Gaps 3;

QY 164 AAATCTAAAGGGGTTTCGTCACATCAAGGAGCTTGGAAACCGGCTGTGGGATGTG 223  
Db 411 AAGCTCAAGGGCTTGGTGTGTGACCCAGCTCCAGGGCTGGGACCCAGGCTGGGACGTG 470

QY 224 GCCGACTTCGTGAAGCCACCTCAGGGAGAGACGTGTCTTCTTGGTGACCAACTTCCTT 283  
Db 471 GCTGACTATGTCTTCCAGGACACAGGGGACAGCTCTCTTTGTAGTTATGACCAACTTCATC 530

QY 284 GTGACGCCAGCCCAAGTTTCAGGGGAGATGCCAGAGCACCGCTCCGTCCTACCTGGCTAAC 343  
Db 531 GTGACCCCTCAGCAGACTCAGGCCATTGTGACAGAAACCC-----AGAAGGTGCATA 584

QY 344 TGCTGGGTGACAGGAGTCCCGGAGGGGAGGGAGGCACACAGCCACCGGTGTAAAAA 403  
Db 585 TGCCAGGATGACAGTGGCTGCACTCCAGGAAAGAGAGAAAGGAAAGCCCAAGGTATTCGC 644

QY 404 ACAGGCCAGTGTGGTGTCTCAATGGGACCCACAGGACCTGTGAGATCTGGAGTTGGTGC 463  
Db 645 ACAGGCAACTGTGTGTCCTTCAATGGCACTGTGAAGACATGTGAGATCTTTGGTTGGTGT 704

QY 464 CCACTGGAGAGTGGCGTTGTGCCCTCGAGGC-----CCTGTGCCCCAGGCCCCAGAACTTC 520  
Db 705 CCTGTAGAGGTGATGACAAGATCCCAAGCCCTGCTCTTCTTCTGAGGCTGAGAACTTC 764

QY 521 ACACCTGTTGATCAAAAACACAGTCACCTTCAGCAAGTTCAACTTCTTCTAAGTCCATGCC 580  
Db 765 ACCCTCTTTCATCAAAAACACAGTCACCTTTCACGCTTTCAGGTCAACAGGCGCAACTCG 824

QY 581 TTGGAGACCTGGGACCCCACTTATTTAAAGCACTGCCGCTATGAACCAACAATTCAGCCGCC 640  
Db 825 GTAGAGGAGGTGAACGGCACCTACATGAAGAAGTGCCTCTATCACAAGATTCAACACCCC 884

QY 641 TACTGTCCCGTGTTCGGCATTTGGGACCTCTGTGGCCCAAGCTGGAGGGACCTTTCAGAGGAC 700  
Db 885 CTGTGCCCACTTCAACCTTGGCTATGTGTGCGAGAGTCAAGCCAGGACTTCGCGCAGC 944

QY 701 CTGGCGTTGCTGGTGGCTCTGTAGGCATCAGAGTTTCACTGGGATTTGTGACCTGGACACC 760  
Db 945 CTGTCTGAGAAGGGTGGGTGGTGTGTAATCACCATTGACTGGAAGTGTGATCTGGACTGG 1004

QY 761 GGGGACTCTGGTGTGGCTCACTACTCTCTCCAGCTGCAGGAGAAAGAGTACAACTTC 820  
Db 1005 CAGCTTCGGCACTGCAAAACCCATCTACCAGTTCCACGGACTGTATGGGGAGAAAGACCTG 1064

QY 821 AGGACAGCCACTCACTGGTGG-----GAGCAACCGGTGTGGAGGCCCGC 865  
Db 1065 TCTCCAGGCTTCAACTTCAGATTTCAGAGTTCGCGCAAGATGGGACAAACCGTCTGT 1124

QY 866 ACCCTGCTCAAGCTCTATGGAATCCGCTTTCGACATCTCTGTCACCGGGCAGGACGGGAG 925  
Db 1125 CACCTCTTCAAGGTGTTGGGATTCACTTTGATATCCTTTGTGATGGCAAGGCTGGGAG 1184

QY 926 TTGGGGCTCATCCCAAGCCCTCACTGGGACCGGGGAGCTTGGCTGGGGCGTGGTC 985  
Db 1185 TTGACATCATCCCTACTATGACTACTATCGGTTCTGGGATTTGGCATCTTTGGAGTGGCC 1244

QY 986 ACCTTTTCTGTGACCTGCTACTGCTGTATGATGAGAGAGCCCACTTCTACTGGAGG 1045  
Db 1245 ACAGTCTTTGTGATCTCTTATTTGCTCCACATCTCTGCTAGAGGCATCTACTACAAGCAG 1304

QY 1046 ACAAGATATGAGGAGGCCAAGGCC 1069  
Db 1305 AAGAAAGTTCAAAATATGCCGAGGAC 1328

Search completed: November 21, 2004, 17:43:04  
Job time : 230.653 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 1290.69 Seconds  
(without alignments)  
10552.804 Million cell updates/sec

Title: US-09-820-095B-1  
Perfect score: 2693  
Sequence: 1 ttgtgactcgtgcccgc.....aaaaaaaaaaaaaaaaaaaaa 2693

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2693	100.0	2693	8 AAL51048	Aal51048 Human P2X
2	1592.6	59.1	16449	8 AAL51049	Aal51049 Human P2X
3	1513.8	56.2	1697	2 AAV61833	Aav61833 Coding se
4	1174	43.6	1452	12 ADP49178	Adp49178 Human P2X
5	1174	43.6	1452	12 ADP49172	Adp49172 Human P2X
6	1174	43.6	1452	12 ADP49182	Adp49182 Human P2X
7	1174	43.6	1452	12 ADP49174	Adp49174 Rat P2X2
8	1174	43.6	1452	12 ADP49176	Adp49176 Human P2X
9	1162.4	43.2	1360	4 AAD04978	Aad04978 Human pur
10	1155.8	42.9	1293	2 AAV61832	Aav61832 Coding se
11	678.4	25.2	800	12 ADP28208	Adp28208 Human sec
12	567.4	21.1	569	4 AAL18520	Aal18520 Probe #84
13	567.4	21.1	569	4 ABA63528	Abas63528 Human foe
14	567.4	21.1	569	4 AAI43639	Aai43639 Probe #12
15	567.4	21.1	569	4 ABA30724	Abas30724 Probe #91
16	567.4	21.1	569	4 AAK37767	Aak37767 Human bon
17	567.4	21.1	569	4 AAK12055	Aak12055 Human bra
18	567.4	21.1	569	4 ABA37406	Abas37406 Human liv
19	567.4	21.1	569	6 ABS11757	Abs11757 Human gen
20	394.4	14.6	396	4 AAD04976	Aad04976 5' end of
21	377	14.0	577	4 AAI19061	Aai19061 Probe #89

ALIGNMENTS

RESULT 1

AAAL51048	AAAL51048 standard; cDNA; 2693 BP.
XX	AC
XX	AAAL51048;
DT	20-FEB-2003 (first entry)
XX	Human P2X-like purinergic receptor G-protein coupled receptor cDNA.
DE	Human; gene; ss; gene therapy; G-protein coupled receptor; chromosome 22;
DE	P2X-like purinergic receptor; anaplastic oligodendroglioma; GPCR;
KW	chronic lymphocytic leukaemia; lung carcinoma; colon carcinoma;
KW	brain carcinoma; drug screening; tissue typing; pharmacogenomic analysis;
KW	single nucleotide polymorphism; SNP.
XX	Hom sapiens.
OS	Hom sapiens.
XX	Key
FT	5'UTR
FT	Location/Qualifiers
FT	1..40
FT	/*tag= a
FT	41..1258
FT	/*tag= b
FT	/product= "Human P2X-like purinergic receptor-related G-
FT	protein coupled receptor"
FT	1259..2693
FT	/*tag= C
FT	3'UTR
FT	WO200279229-A2.
FT	10-OCT-2002.
XX	28-MAR-2002; 2002WO-US009545.
XX	29-MAR-2001; 2001US-00820095.
XX	(PEKE ) PE CORP NY.
XX	Wei M, Gong F, Di Francesco V, Beasley EM;
XX	WPI; 2003-040648/03.
XX	P-PSDB; AAO15997.
XX	New peptides related to P2X-like purinergic receptor subfamily, useful for



```
OY 1861 CCCAGGAAGCGCGAGCATGCTCTTGGGCCACACCTGCTAGTTTATGAGACCGGC 1920
Dd 1861 CCCAGGAAGCGCGAGCATGCTCTTGGGCCACACCTGCTAGTTTATGAGACCGGC 1920
OY 1921 TGCCTTCCAGGTGAGCGCTTTTCCCATGAGAGTCTCGGAGAGAGACAGAGGCGGCAG 1980
Dd 1921 TGCCTTCCAGGTGAGCGCTTTTCCCATGAGAGTCTCGGAGAGAGACAGAGGCGGCAG 1980
OY 1981 GGCCTAAGTGTGATCATCTGGTCTTCAGACCTCTATATCCCTCGTAAACCC 2040
Dd 1981 GGCCTAAGTGTGATCATCTGGTCTTCAGACCTCTATATCCCTCGTAAACCC 2040
OY 2041 CAGCCCAACCCCTTGGAACTCTTCTCAGGCTTCCCTCAGAGCCCTGGGGTGGAGGCT 2100
Dd 2041 CAGCCCAACCCCTTGGAACTCTTCTCAGGCTTCCCTCAGAGCCCTGGGGTGGAGGCT 2100
OY 2101 GTGGAGGCTGTATCATCTGAATTCATCTAGTCCAACTATCTAGTAAAGCTGTCTGG 2160
Dd 2101 GTGGAGGCTGTATCATCTGAATTCATCTAGTCCAACTATCTAGTAAAGCTGTCTGG 2160
OY 2161 GCAGCTGCTCAGGAGGCGCTGCTCTGTATCCAGGCTGGATGGCTGGGAAGAA 2220
Dd 2161 GCAGCTGCTCAGGAGGCGCTGCTCTGTATCCAGGCTGGATGGCTGGGAAGAA 2220
OY 2221 TGGTTCACAAACACACACCGAGATCTCCCTCAGGCTGGCCAGGTTTGCAGCTGGAAT 2280
Dd 2221 TGGTTCACAAACACACACCGAGATCTCCCTCAGGCTGGCCAGGTTTGCAGCTGGAAT 2280
OY 2281 CTCCTCTGTGCTCCAGGCGGCGAGGAATCTAAGTGTCCACCCAGGAGGCAAGGG 2340
Dd 2281 CTCCTCTGTGCTCCAGGCGGCGAGGAATCTAAGTGTCCACCCAGGAGGCAAGGG 2340
OY 2341 GCTGCTTCCACTGTGGTACTCTGGTATCAGGCAAGCTGTGAGGCGCCAGGGGTGGG 2400
Dd 2341 GCTGCTTCCACTGTGGTACTCTGGTATCAGGCAAGCTGTGAGGCGCCAGGGGTGGG 2400
OY 2401 CTGAGACTGGCTGACATCTAGATCAGCTGCCACCTGGAGCCTCAGTAAATGCCCTGG 2460
Dd 2401 CTGAGACTGGCTGACATCTAGATCAGCTGCCACCTGGAGCCTCAGTAAATGCCCTGG 2460
OY 2461 GTCCCTGTGCTCTCAATCTCCAGGACCATGTCCATGGGAGGTGGGCTCTGAAGGGCG 2520
Dd 2461 GTCCCTGTGCTCTCAATCTCCAGGACCATGTCCATGGGAGGTGGGCTCTGAAGGGCG 2520
OY 2521 AAGTGGAGAGCAGGCGCCCTGAGCCTGGGTATCCAAAGAGGGGACGTCACCTGAT 2580
Dd 2521 AAGTGGAGAGCAGGCGCCCTGAGCCTGGGTATCCAAAGAGGGGACGTCACCTGAT 2580
OY 2581 TCTCCTTGGGCCACAGGAGAGCTGATGCTATGCTGACAAAGTCAAGGAGTAAAGCCA 2640
Dd 2581 TCTCCTTGGGCCACAGGAGAGCTGATGCTATGCTGACAAAGTCAAGGAGTAAAGCCA 2640
OY 2641 GCAAGGCCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2693
Dd 2641 GCAAGGCCACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2693

RESULT 2
AAL51049
ID AAL51049 standard; DNA; 16449 BP.
AC AAL51049;
XX
XX
DT 20-FEB-2003 (first entry)
DE Human p2X-like purinergic receptor G-protein coupled receptor gene.
XX
XX Human; gene; ds; gene therapy; G-protein coupled receptor; chromosome 22;
XX p2X-like purinergic receptor; anaplastic oligodendroglioma; GPCR;
XX chronic lymphocytic leukaemia; lung carcinoma; colon carcinoma;
XX brain carcinoma; drug screening; tissue typing; pharmacogenomic analysis;
XX SNP; single nucleotide polymorphism.
XX
```

FT intron 9871..9952  
 FT /tag= x  
 FT /number= 5  
 FT 9953..10033  
 FT /tag= y  
 FT /number= 6  
 FT 10034..10109  
 FT /tag= z  
 FT /number= 6  
 FT 10110..10251  
 FT /tag= aa  
 FT /number= 7  
 FT 10252..10343  
 FT /tag= ab  
 FT /number= 7  
 FT 10344..10453  
 FT /tag= ac  
 FT /number= 8  
 FT 10454..12638  
 FT /tag= ad  
 FT /number= 8  
 FT variation replace(10504,C)  
 FT /tag= ae  
 FT /note= "Single nucleotide polymorphism"  
 FT replace(12609,A)  
 FT /tag= af  
 FT /note= "Single nucleotide polymorphism"  
 FT 12639..12732  
 FT /tag= ag  
 FT /number= 9  
 FT 12733..12844  
 FT /tag= ah  
 FT /number= 9  
 FT 12845..12910  
 FT /tag= ai  
 FT /number= 10  
 FT 12911..13085  
 FT /tag= aj  
 FT /number= 10  
 FT 13086..13163  
 FT /tag= ak  
 FT /number= 11  
 FT 13164..13253  
 FT /tag= al  
 FT /number= 11  
 FT 13254..13451  
 FT /tag= am  
 FT /number= 12  
 FT variation replace(13367,A)  
 FT /tag= an  
 FT /note= "Single nucleotide polymorphism"  
 FT replace(14191,C)  
 FT /tag= ao  
 FT /note= "Single nucleotide polymorphism"  
 FT replace(14227,G)  
 FT /tag= ap  
 FT /note= "Single nucleotide polymorphism"  
 FT variation replace(15027,C)  
 FT /tag= aq  
 FT /note= "Single nucleotide polymorphism"  
 FT replace(15441,C)  
 FT /tag= ar  
 FT /note= "Single nucleotide polymorphism"  
 FT  
 XX W0200279229-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 28-MAR-2002; 2002MO-US009545.  
 XX  
 PR 29-MAR-2001; 2001US-00820095.  
 XX  
 PA (PEKE ) PE CORP NY.

XX  
 PI Wei M, Gong F, Di Francesco V, Beasley EM;  
 XX WPI; 2003-040648/03.  
 DR P-PSDB; AAO15997.  
 XX  
 PT New peptides related to P2X-like purigenic receptor subfamily, useful for  
 PT treating disorders associated with abnormal expression of protease in  
 PT anaplastic oligodendroglioma, leukemia, carcinoid lung, or large cell  
 PT lung carcinoma.  
 XX  
 PS Claim 4; Fig 3A-E; 87pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequence (located on  
 CC chromosome 22) of a human G-protein coupled receptor (GPCR) which is  
 CC related to the P2X-like purigenic receptor subfamily. The DNA and  
 CC protein sequences of the invention are useful for treating: anaplastic  
 CC oligodendroglioma; chronic lymphocytic leukaemia; lung carcinoma; colon  
 CC carcinoma; and brain carcinoma. The DNA and protein sequences of the  
 CC invention are also useful for drug screening assays, tissue typing and  
 CC pharmacogenomic analysis. The present genomic DNA sequence represents the  
 CC gene encoding the human GPCR that is related to the P2X-like purinergic  
 CC receptor subfamily  
 XX  
 SQ Sequence 16449 BP; 3284 A; 4727 C; 4589 G; 3849 T; 0 U; 0 Other;  
 Query Match 59.1%; Score 1592.6; DB 8; Length 16449;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1595; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1053 ATGAGGAGCCCAAGGCCCGGAAAGCAACCGCCAACTCTGTGTGGAGGAGTGGCCCTTG 1112  
 DB 13246 ATCTGAGGCCAAGGCCCGGAAAGCAACCGCCAACTCTGTGTGGAGGAGTGGCCCTTG 13305  
 QY 1113 GATCCCAAGCCGACTGGCGGAGTGCCTCAGACGGAGCTCAGACCTGCACCCACGGCCA 1172  
 DB 13306 CATCCCAAGCCGACTGGCGGAGTGCCTCAGACGGAGCTCAGACCTGCACCCACGGCCA 13365  
 QY 1173 CTGCTGCTGGGAGTCAGACACAGACACCCAGATGGCCCTGTCCAAAGTTCTGACACCCACT 1232  
 DB 13366 CTGCTGCTGGGAGTCAGACACAGACACCCAGATGGCCCTGTCCAAAGTTCTGACACCCACT 13425  
 QY 1233 TGCCAAACCATTCGGGAGCCTGTAGCGTTCCCTGTGTGTGTGAGTGTGGGCTGGGA 1292  
 DB 13426 TGCCAAACCATTCGGGAGCCTGTAGCGTTCCCTGTGTGTGTGAGTGTGGGCTGGGA 13485  
 QY 1293 AGGGCGGGCCCTGCCTGGGGATCTCAAGGATGAGGCCCGCAGCATGGAGGATTCGGGGTA 1352  
 DB 13486 AGGGCGGGCCCTGCCTGGGGATCTCAAGGATGAGGCCCGCAGCATGGAGGATTCGGGGTA 13545  
 QY 1353 GAATTCACCCCTTGAACCCCGCAGACAGTCCTCCCTCCCTGACTCCACCTTTGGTAGGGTG 1412  
 DB 13546 GAATTCACCCCTTGAACCCCGCAGACAGTCCTCCCTCCCTGACTCCACCTTTGGTAGGGTG 13605  
 QY 1413 CTGCCTCAGGAGCCATAGAGTCGGCTGTGTTTGTAGACGGCGCAGAACCTGACCCGT 1472  
 DB 13606 CTGCCTCAGGAGCCATAGAGTCGGCTGTGTTTGTAGACGGCGCAGAACCTGACCCGT 13665  
 QY 1473 GGAGACTGGGAGAGCCCGCAGGACCTGTATTTCAGGGGTCCGACTGCGATGTGGCAGGG 1532  
 DB 13666 GGAGACTGGGAGAGCCCGCAGGACCTGTATTTCAGGGGTCCGACTGCGATGTGGCAGGG 13725  
 QY 1533 GCTCTGCTGGCTGTGGGCTGGAGGTCTCTCTCCAGTGCCTGTGTCTCCAGTGTTCCTA 1592  
 DB 13726 GCTCTGCTGGCTGTGGGCTGGAGGTCTCTCTCCAGTGCCTGTGTCTCCAGTGTTCCTA 13785  
 QY 1593 GCAGAGGTATGCTTACCAGTGTGCAGCAGACAGCCCTCTGCTGGGTCTTGGCCCTC 1652  
 DB 13786 GCAGAGGTATGCTTACCAGTGTGCAGCAGACAGCCCTCTGCTGGGTCTTGGCCCTC 13845  
 QY 1653 CTCCCCCATCTGCACCCCCCATCATAGGTAGAGACCCACCCCTCCCATCGGTCTCATG 1712  
 DB 13846 CTCCCCCATCTGCACCCCCCATCATAGGTAGAGACCCACCCCTCCCATCGGTCTCATG 13905

QY 1713 GGCTGTGCAGCTGGAGCCAAAGGCAAGCAGAAAGAGAGTGTATGGGGAGGGGATT 1772  
 DE |||||  
 XX |||||  
 Db 13906 GGCTGTGCAGCTGGAGCCAAAGGCAAGGTAGAAAGAGAGTGTATGGGGAGGGGATT 13965  
 QY 1773 GTTTCAGCTTCTGTGTGTGTATGCCCCAGAGAGTCTTAATCTAGGAAATGGGGTGG 1832  
 Db 13966 GTTTCAGCTTCTGTGTGTGTATGCCCCAGAGAGTCTTAATCTAGGAAATGGGGTGG 14025  
 QY 1833 AGTAGGAGATATACACCTCCCTATCCCCAGGCAAGGCGAGAGATGTCTTGGGGC 1892  
 Db 14026 AGTAGGAGATATACACCTCCCTATCCCCAGGCAAGGCGAGAGATGTCTTGGGGC 14085  
 QY 1893 CACACCTGTATTAGTTATCAGGACCGGCTCTTCCAGTGGTAGCCCTTTTCCCATGGAG 1952  
 Db 14086 CACACCTGTATTAGTTATCAGGACCGGCTCTTCCAGTGGTAGCCCTTTTCCCATGGAG 14145  
 QY 1953 GTCTGGGAGAGAGAGAGAGGCGGCGAGGCTTAAGTTGGTGATCAATGGGTTCTTCAGGA 2012  
 Db 14146 GTCTGGGAGAGAGAGAGAGGCGGCGAGGCTTAAGTTGGTGATCAATGGGTTCTTCAGGA 14205  
 QY 2013 CTTTCTATATCCCTCTCGGTAAACCCCCAGCCCAACCCCTTGGAACTTTTCTCCAGGC 2072  
 Db 14206 CTTTCTATATCCCTCTCGGTAAACCCCCAGCCCAACCCCTTGGAACTTTTCTCCAGGC 14265  
 QY 2073 TTCTGAGAGCCCTGGGGTGGAGAGCTGTGGAGGCTGTACCTGAATTCATTCAAG 2132  
 Db 14266 TTCTGAGAGCCCTGGGGTGGAGAGCTGTGGAGGCTGTACCTGAATTCATTCAAG 14325  
 QY 2133 TCCAACTATACCTAGGAAGCTGTCTGGGACGCTGCTGAGGAGGCGCTGCTGATC 2192  
 Db 14326 TCCAACTATACCTAGGAAGCTGTCTGGGACGCTGCTGAGGAGGCGCTGCTGATC 14385  
 QY 2193 CCAGGCTGATGGAGTGGTGAAGAAATGTTCCAAACAAACACCCAGAGATCTCCCTC 2252  
 Db 14386 CCAGGCTGATGGAGTGGTGAAGAAATGTTCCAAACAAACACCCAGAGATCTCCCTC 14445  
 QY 2253 AGGCTGGCCAGGTTTTCAGCTGGAATCTCTCTGTTCCAGGCGGGGAGGGAATT 2312  
 Db 14446 AGGCTGGCCAGGTTTTCAGCTGGAATCTCTCTGTTCCAGGCGGGGAGGGAATT 14505  
 QY 2313 CTAAGTGTCCACCCAGGAGGCAAGGGGCTGCTTCCACTGTGGGTACCTGGTGATCAG 2372  
 Db 14506 CTAAGTGTCCACCCAGGAGGCAAGGGGCTGCTTCCACTGTGGGTACCTGGTGATCAG 14565  
 QY 2373 GGCAGAGCTGTGAGGCGCCAGGGTGGGCTGAGACTGGGCTGACATCTAGAAATCACCTGC 2432  
 Db 14566 GGCAGAGCTGTGAGGCGCCAGGGTGGGCTGAGACTGGGCTGACATCTAGAAATCACCTGC 14625  
 QY 2433 CACCTGGAGCCTCAGTAAATGCTTGGGTCCCTGCTCCTCAATCTCCAGAGCCATG 2492  
 Db 14626 CACCTGGAGCCTCAGTAAATGCTTGGGTCCCTGCTCCTCAATCTCCAGAGCCATG 14685  
 QY 2493 TCCATGGGAGTGGGCTCTGAAGGGCGAAGGTGGAGAGCAGAGGCCCTCAGGCTGGG 2552  
 Db 14686 TCCATGGGAGTGGGCTCTGAAGGGCGAAGGTGGAGAGCAGAGGCCCTCAGGCTGGG 14745  
 QY 2553 TATCAAGAGGGGACGCTGACCTGATTCTCTTGGGGCCAGAGGAGCTGATCAT 2612  
 Db 14746 TATCAAGAGGGGACGCTGACCTGATTCTCTTGGGGCCAGAGGAGCTGATCAT 14805  
 QY 2613 GGCTGGACAAAGTCAAGGAGTAAAGCCAGCAAGGCCACC 2651  
 Db 14806 GGCTGGACAAAGTCAAGGAGTAAAGCCAGCAAGGCCACC 14844

## RESULT 3

AAV61833

ID AAV61833 standard; cDNA; 1697 BP.

XX

XX AAV61833;

XX

DT 28-APR-1999 (first entry)

XX Coding sequence for human p53 regulated protein, P2XM.  
 DE  
 XX  
 KW Human; p53 regulated protein; P2XM; ATP receptor; RP-2 protein;  
 KW thymocyte; apoptosis; tumour-suppressor gene; rhabdoid sarcoma;  
 KW anticancer drug; ss.  
 XX  
 OS Homo sapiens.

PH Key Location/Qualifiers  
 FT CDS 46..1341  
 FT /\*tag= a

XX WO9842835-A1.

XX 01-OCT-1998.

XX 18-MAR-1998; 98WO-JP001146.

XX 26-MAR-1997; 97JP-00093044.

XX (SAKA ) OTSUKA PHARM CO LTD.

XX Tokino T, Nakamura Y;

XX WPI; 1998-532006/45.

XX P-PSDB; AAW76434.

Human gene P2XM whose transcription is induced by p53 - useful, e.g. for  
 diagnostic purposes and in development of new anticancer drugs.

XX Disclosure; Page 23-27; 43pp; Japanese.

This sequence represents the coding sequence for the P2XM protein of the  
 invention. The protein is significantly homologous to: (i) the P2X family  
 of ATP receptors, and (ii) RP-2 protein which is expressed in thymocytes  
 during apoptosis. Transcription of the genes is specifically regulated by  
 the tumour-suppressor gene p53. The P2XM gene is specifically expressed  
 in skeletal muscle and has been localised to chromosome 22q11, an area  
 where mutation and sequence losses frequently occur in rhabdoid sarcomas.  
 The genes may be used for diagnostic purposes (e.g. by detecting changes  
 occurring in the gene in sarcomas), using probes and primers containing  
 or derived from all or part of the genes. The genes may further be used  
 in the development of new anticancer drugs

XX SQ Sequence 1697 BP; 343 A; 496 C; 520 G; 338 T; 0 U; 0 Other;

Query Match 56.2%; Score 1513.8; DB 2; Length 1697;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1511; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 97 GTGGGCTCTCTCGCCAAAAGGCTACCGAGGCGGACCTGGAACCCAGTTTCCAT 156

Db 180 GTGGGCGCTCTCGCCAAAAGGCTACCGAGGCGGACCTGGAACCCAGTTTCCAT 239

QY 157 CATCACCAACTCAAGGGGTTTCCCTCACTCAGTCAAGGAGCTTGGAAACCGGCTGTG 216

Db 240 CATCACCAACTCAAGGGGTTTCCCTCACTCAGTCAAGGAGCTTGGAAACCGGCTGTG 299

QY 217 GGATGTGGCGGCTTCGTGAAGCCACTCAGGGAGAGAACTGTTCTTCTTGTGACCAA 276

Db 300 GGATGTGGCGGCTTCGTGAAGCCACTCAGGGAGAGAACTGTTCTTCTTGTGACCAA 359

QY 277 CTTTCTTGTGACGCCAGCCCAAGTTTCAGGGCAGATGCCAGACCCCGTCCCACT 336

Db 360 CTTTCTTGTGACGCCAGCCCAAGTTTCAGGGCAGATGCCAGACCCCGTCCCACT 419

QY 337 GGCTAACTGTGGGTCCAGGAGACTGCCCCGAAGGGAGGGAGGCACACACAGCCACGG 396

Db 420 GGCTAACTGTGGGTCCAGGAGACTGCCCCGAAGGGAGGGAGGCACACACAGCCACGG 479

QY 397 TGTAAAAACAGGCCAGTGTGTGTGTTCAATGGGACCCACAGGACCTGTGAGATCTGGAG 456







QY 277 CTTCTCTGTGAGCCAGCCCAAGTTTCAGGGGAGATGCCCCAGAGACCCCGTCCGTCCCACT 336  
Db 359 CTTCTCTGTGAGCCAGCCCAAGTTTCAGGGGAGATGCCCCAGAGACCCCGTCCGTCCCACT 418  
QY 337 GGTAACTCTGGTTCGAGAGAGTCCGCCGAAGGGAGGAGGACACACAGCCACGG 396  
Db 419 GGTAACTCTGGTTCGAGAGAGTCCGCCGAAGGGAGGAGGACACACAGCCACGG 478  
QY 397 TGTAAAAACAGGCCAGTGTGTGTGTGTTCATATGGGACCCACAGGACCTGTGAGATCTGAG 456  
Db 479 TGTAAAAACAGGCCAGTGTGTGTGTGTTCATATGGGACCCACAGGACCTGTGAGATCTGAG 538  
QY 457 TTGGTCCAGTGGAGAGTGGGTGTGGCCCTCGAGGCCCTGTGGGCCAGGCCAGAA 516  
Db 539 TTGGTCCAGTGGAGAGTGGGTGTGGCCCTCGAGGCCCTGTGGGCCAGGCCAGAA 598  
QY 517 CTTACACACTGTTTCATCAAAAAACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTCCAA 576  
Db 599 CTTACACACTGTTTCATCAAAAAACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTCCAA 658  
QY 577 TGCCTTGGAGACCTGGGACCCACCTATTTTAAGCACTGCCCTATGAACCAATTCAG 636  
Db 659 TGCCTTGGAGACCTGGGACCCACCTATTTTAAGCACTGCCCTATGAACCAATTCAG 718  
QY 637 CCGCTACTGTCCGTGTTCCGCAATTGGGACCTCTGTGGCCAAAGGCTGGAGGACCTTCGA 696  
Db 719 CCGCTACTGTCCGTGTTCCGCAATTGGGACCTCTGTGGCCAAAGGCTGGAGGACCTTCGA 778  
QY 697 GGAACCTGGCGTGTCTGGTGTGTGTAGGCATCAAGTTCACTGGATTTGACCTGGA 756  
Db 779 GGAACCTGGCGTGTCTGGTGTGTGTAGGCATCAAGTTCACTGGATTTGACCTGGA 838  
QY 757 CACCGGGACTGTGGCTGTGGCTCACTACTCTTCCAGCTGCAGGAGAGAGCTACAA 816  
Db 839 CACCGGGACTGTGGCTGTGGCTCACTACTCTTCCAGCTGCAGGAGAGAGCTACAA 898  
QY 817 CTTAGGACAGCACTCACTGTGGGAGCAACCGGCTGTGGAGGCCCGCACCTCTCTCAA 876  
Db 899 CTTAGGACAGCACTCACTGTGGGAGCAACCGGCTGTGGAGGCCCGCACCTCTCTCAA 958  
QY 877 GCTCTATGAATCCGCTTGCATCTGTCTACCGGGACGGCAGGAGAGTTCCGGCTCAT 936  
Db 959 GCTCTATGAATCCGCTTGCATCTGTCTACCGGGACGGCAGGAGAGTTCCGGCTCAT 1018  
QY 937 CCCCAGCGCGTCACTGGGACACCGGGGACGCTTGGCTGGCGTGGTCACTTTTCTG 996  
Db 1019 CCCCAGCGCGTCACTGGGACACCGGGGACGCTTGGCTGGCGTGGTCACTTTTCTG 1078  
QY 997 TGACCTGTCTACTGTGTATGTGGATAGAGAACCCATTTCTACTGGAGACAAAGTATGA 1056  
Db 1079 TGACCTGTCTACTGTGTATGTGGATAGAGAACCCATTTCTACTGGAGACAAAGTATGA 1138  
QY 1057 GGAGGSCAAGGCCCGGAAAGAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1116  
Db 1139 GGAGGSCAAGGCCCGGAAAGAACCGCCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1198  
QY 1117 CCAAGCCGACTGGCGAGTGTCTCAGACGGAGCTCAGCACTGCACCCACCGCCACTGC 1176  
Db 1199 CCAAGCCGACTGGCGAGTGTCTCAGACGGAGCTCAGCACTGCACCCACCGCCACTGC 1258  
QY 1177 TGCTGGAGTTCAGACACAGACACAGGATGGCCCTGTCCAAGTTCTGACACCCACTTGC 1236  
Db 1259 TGCTGGAGTTCAGACACAGACACAGGATGGCCCTGTCCAAGTTCTGACACCCACTTGC 1318  
QY 1237 AACCCATTCCGGAGGCTGTAGCCG-TTCCCTGTCTGGTTGAGA-----GTTGGGGG 1286  
Db 1319 AACCCATTCCGGAGGCTGTAGCCGTTTCCCTGTCTGGTTGAGAAGAGAGAGGGGCTGGG 1378  
QY 1287 CTGGGAAGGGCGGGCCCTGTCTGGGATCTCAAGATAGAGCCCCAGATGA 1340  
Db 1379 AAGGAAGGAGCCCTGTCCCTGCGGAGCGAAAGCAAGGATGAGGCAACAGCAATGA 1432

RESULT 5  
ADP49172  
ID ADP49172 standard; DNA; 1452 BP.  
XX  
AC ADP49172;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
Human P2X1 purinergic receptor DNA sequence for odour modulation.  
DE ds; Gene; odour sensitivity; P2X purinergic receptor;  
XX P2Y purinergic receptor; signal transduction pathway;  
KW olfactory signalling; micro-array.  
XX  
OS Homo sapiens.  
XX  
PN WO2004047749-A2.  
XX  
PD 10-JUN-2004.  
XX  
PF 21-NOV-2003; 2003WO-US037389.  
XX  
PR 21-NOV-2002; 2002US-0428140P.  
XX  
(UTAH ) UNIV UTAH RES FOUND.  
XX  
PI Lucero M, Hegg C;  
XX  
DR WPI; 2004-460642/43.  
XX  
Modulating odor sensitivity in a subject, comprises administering a  
PT composition comprising an agonist or antagonist of P2X or P2Y purinergic  
PT receptor to the subject.  
XX  
PS Disclosure; SEQ ID NO 1; 108pp; English.  
XX  
The invention relates to a method of modulating (M1) odour sensitivity in  
CC a subject, by administering a composition which is an agonist or  
CC antagonist of a P2X or P2Y purinergic receptor. (M1) is useful for  
CC modulating odour sensitivity in a subject (claimed). The compositions  
CC used for modulating odour sensitivity in a subject are useful for  
CC studying the signal transduction pathways related to olfactory signaling.  
CC The compositions are also useful as reagents in micro-arrays or as  
CC reagents to probe or analyze existing micro-arrays. This sequence  
CC corresponds to the human P2X1 DNA sequence.  
XX  
SQ Sequence 1452 BP; 316 A; 415 C; 442 G; 279 T; 0 U; 0 Other;

Query Match 43.6%; Score 1174; DB 12; Length 1452;  
Best Local Similarity 97.2%; Pred. No. 3.2e-267;  
Matches 1219; Conservative 0; Mismatches 25; Indels 10; Gaps 2;  
QY 97 GTGGGCTCTCTCGCCAAAAGGCTACAGGAGCGGACCTGGAACCCCGATTTTCCAT 156  
Db 179 GTGGGCTCTCTCGCCAAAAGGCTACAGGAGCGGACCTGGAACCCCGATTTTCCAT 238  
QY 157 CATCACCAAAACCTCAAGGGGTTTCCGTCACTCAGATCAAGGAGCTTGGAAACCGGCTGTG 216  
Db 239 CATCACCAAAACCTCAAGGGGTTTCCGTCACTCAGATCAAGGAGCTTGGAAACCGGCTGTG 298  
QY 217 GGATGTGGCGACTTGTGTGAAGCCACTCAGGAGAGAGAGCTGTCTTCTTGTGTGACCA 276  
Db 299 GGATGTGGCGACTTGTGTGAAGCCACTCAGGAGAGAGAGCTGTCTTCTTGTGTGACCA 358  
QY 277 CTTCTCTGTGACGCCACCCCAAGTTTTCAGGGCAGATGCCAGAGACCCCGTCCGCTCCACT 336  
Db 359 CTTCTCTGTGACGCCACCCCAAGTTTTCAGGGCAGATGCCAGAGACCCCGTCCGCTCCACT 418  
QY 337 GGCTAACTGTGGTTCGAGGAGCTGCCCGGAAGGGAGGAGGACACACAGACCAACGG 396  
Db 419 GGCTAACTGTGGTTCGAGGAGCTGCCCGGAAGGGAGGAGGACACACAGACCAACGG 478

QY 397 TGTAAACAGCCAGTGTGGTGTTCATATGGACCCACAGGACCTGTGAGATCTGGAG 456  
 Db 479 TGTAAACAGCCAGTGTGGTGTTCATATGGACCCACAGGACCTGTGAGATCTGGAG 538  
 QY 457 TTGGTGCCAGTGGAGAGTGGCGTTGTGCCCTCGAGGCCCTGTGGCCAGGCCAGAA 516  
 Db 539 TTGGTGCCAGTGGAGAGTGGCGTTGTGCCCTCGAGGCCCTGTGGCCAGGCCAGAA 598  
 QY 517 CTTACACATGTTTCATCAAAACACAGTCACTTCAAGCAAGTTCACATCTCTAGTCCAA 576  
 Db 599 CTTACACATGTTTCATCAAAACACAGTCACTTCAAGCAAGTTCACATCTCTAGTCCAA 658  
 QY 577 TGCCTTGGAGACCTGGGACCCCACTATTTAAGCACTGCGCTATGAACCACAATTCAG 636  
 Db 659 TGCCTTGGAGACCTGGGACCCCACTATTTAAGCACTGCGCTATGAACCACAATTCAG 718  
 QY 637 CCCCTACTGTCCTGGTTCGCAATGGGACCTCGTGCCCAAGCTGGAGGACCTTCGA 696  
 Db 719 CCCCTACTGTCCTGGTTCGCAATGGGACCTCGTGCCCAAGCTGGAGGACCTTCGA 778  
 QY 697 GGACCTGCGTGTGCTGGGTGCTCTGTAGGCATCAGAGTTCACCTGGGATTTGACCTGGA 756  
 Db 779 GGACCTGCGTGTGCTGGGTGCTCTGTAGGCATCAGAGTTCACCTGGGATTTGACCTGGA 838  
 QY 757 CACCGGGACTGTGCTGCTGCTCACTACTCTTCAGCTGCAGAGAGAGCTACAA 816  
 Db 839 CACCGGGACTGTGCTGCTGCTCACTACTCTTCAGCTGCAGAGAGAGCTACAA 898  
 QY 817 CTTAGGACAGCACTCACTGCTGGGAGCAACCGGGTGTGGAGCCCGCACTTCCTCAAA 876  
 Db 899 CTTAGGACAGCACTCACTGCTGGGAGCAACCGGGTGTGGAGCCCGCACTTCCTCAAA 958  
 QY 877 GCTATATGAATTCGCTTCGACATCTCGTCAACCGGAGAGAGAGTTCGGGCTCAT 936  
 Db 959 GCTATATGAATTCGCTTCGACATCTCGTCAACCGGAGAGAGTTCGGGCTCAT 1018  
 QY 937 CCCACGCGCTCACTACCTGGGAGCCGCGGAGCTTGGCTGGGCTGTCACTTTTCTG 996  
 Db 1019 CCCACGCGCTCACTACCTGGGAGCCGCGGAGCTTGGCTGGGCTGTCACTTTTCTG 1078  
 QY 997 TGACCTGCTACTGCTGTATGTGATAGAGAGCCCACTTCTACTGGAGGACAAAGTATGA 1056  
 Db 1079 TGACCTGCTACTGCTGTATGTGATAGAGAGCCCACTTCTACTGGAGGACAAAGTATGA 1138  
 QY 1057 GGAGGCCAAGCCCGAGAGCAACCGGCACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1116  
 Db 1139 GGAGGCCAAGCCCGAGAGCAACCGGCACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1198  
 QY 1117 CCAAGCCGAGCTGGCCGAGTGCCTCAGACGGAGCTCAGCACTGCACCCCAAGCCCACTGC 1176  
 Db 1199 CCAAGCCGAGCTGGCCGAGTGCCTCAGACGGAGCTCAGCACTGCACCCCAAGCCCACTGC 1258  
 QY 1177 TGCTGGGAGTGCAGACACAGAGATGGCCCTGTCCAAAGTTCCTGACACCCCACTGGCC 1236  
 Db 1259 TGCTGGGAGTGCAGACACAGAGATGGCCCTGTCCAAAGTTCCTGACACCCCACTGGCC 1318  
 QY 1237 AACCCATTCCGGGAGCCCTGAGCCG-TTCCCTGTGTGTGAGA-----GTTGGGGG 1286  
 Db 1319 AACCCATTCCGGGAGCCCTGAGCCGTTCCCTGTGTGTGAGAGAGAGGGGCTGGG 1378  
 QY 1287 CTGGGAAGGGGGGGCCCTGCTGGGATCTCAAGGATGAGGCCCCCAAGCATGGA 1340  
 Db 1379 AAGGAAGGACCCCTGCCCTGCCGAGCGAAGCAAGGATGAGGCAACAGCAATGA 1432

## RESULT 6

ADP49182

ID ADP49182 standard; DNA; 1452 BP.

XX AC

XX ADP49182;

XX XX

DT 26-AUG-2004 (first entry)

XX XX

DE Human P2X6 purinergic receptor DNA sequence for odour modulation.  
 XX ds: gene; odour sensitivity; P2X purinergic receptor;  
 KW P2X purinergic receptor; signal transduction pathway;  
 KW olfactory signalling; micro-array.  
 XX Homo sapiens.  
 XX WO2004047749-A2.  
 PN 10-JUN-2004.  
 XX 21-NOV-2003; 2003WO-US037389.  
 XX 21-NOV-2002; 2002US-0428140P.  
 XX (UTAH ) UNIV UTAH RES FOUND.  
 PA Lucero M, Hegg C;  
 PI WPI; 2004-460642/43.  
 DR Modulating odor sensitivity in a subject, comprises administering a  
 PT composition comprising an agonist or antagonist of P2X or P2Y purinergic  
 PT receptor to the subject.  
 XX Disclosure; SEQ ID NO 11; 108pp; English.  
 XX The invention relates to a method of modulating (M1) odour sensitivity in  
 CC a subject, by administering a composition which is an agonist or  
 CC antagonist of a P2X or P2Y purinergic receptor. (M1) is useful for  
 CC modulating odour sensitivity in a subject (claimed). The compositions  
 CC used for modulating odour sensitivity in a subject are useful for  
 CC studying the signal transduction pathways related to olfactory signaling.  
 CC The compositions are also useful as reagents in micro-arrays or as  
 CC reagents to probe or analyze existing micro-arrays. This sequence  
 CC corresponds to the human P2X6 DNA sequence.  
 XX Sequence 1452 BP; 316 A; 415 C; 442 G; 279 T; 0 U; 0 Other;

Query Match 43.6%; Score 1174; DB 12; Length 1452;  
 Best Local Similarity 97.2%; Pred. No. 3.2e-267;  
 Matches 1219; Conservative 0; Mismatches 25; Indels 10; Gaps 2;  
 QY 97 GTGGGCTCTCTCCGCAAAAGAGGCTACAGAGCGGACCTGGAACCCCAAGTTTCCAT 156  
 Db 179 GTGGGCTCTCTCCGCAAAAGAGGCTACAGAGCGGACCTGGAACCCCAAGTTTCCAT 238  
 QY 157 CATCACCAAACTCAAAGGGGTTTCGTCACCTCAGATCAAGGAGCTTGAAAACCGGCTGTG 216  
 Db 239 CATCACCAAACTCAAAGGGGTTTCGTCACCTCAGATCAAGGAGCTTGAAAACCGGCTGTG 298  
 QY 217 GATGTGGCCGACTTCGTGAAGCCACCTCAGGAGAGAACGTGTCTCTTGGTGACAA 276  
 Db 299 GATGTGGCCGACTTCGTGAAGCCACCTCAGGAGAGAACGTGTCTCTTGGTGACAA 358  
 QY 277 CTTCTCTGTAGCCAGCCCAAGTTTCAGGGCAGATGCCAGAGCACCCGTCCTGCCACT 336  
 Db 359 CTTCTCTGTAGCCAGCCCAAGTTTCAGGGCAGATGCCAGAGCACCCGTCCTGCCACT 418  
 QY 337 GCGTAACTGCTGGGTGCGAGGAGACTGCCCGAAGGGAGGAGGACACACAGCCACGG 396  
 Db 419 GCGTAACTGCTGGGTGCGAGGAGACTGCCCGAAGGGAGGAGGACACACAGCCACGG 478  
 QY 397 TGTAAAACACAGCCCAAGTGTGGTGTTCATGGGACCCACAGGACCTGTGAGATCTGGAG 456  
 Db 479 TGTAAAACACAGCCCAAGTGTGGTGTTCATGGGACCCACAGGACCTGTGAGATCTGGAG 538  
 QY 457 TTGGTGCCCAAGTGGAGAGTGGCGTTGTGCCCTCGAGGCCCCCTGTGGCCCCAGGCCAGAA 516  
 Db 539 TTGGTGCCCAAGTGGAGAGTGGCGTTGTGCCCTCGAGGCCCCCTGTGGCCCCAGGCCAGAA 598  
 QY 517 CTTCACTGTGTCACTCAAAACACAGTCACTTCAGCAAGTTCACATCTCTAAGTCCAA 576

Db 599 CTTACACATGTTTCATCAAAAAACAGTACCTTCAGCAAGTTCAACTTCTCTAGTCCAA 658  
 Qy 577 TGCCTTGGAGACTGGGACCCACACCTATTATTAAGCACTGCCGCTATGAACCAAAATTCAG 636  
 Db 659 TGCCTTGGAGACTGGGACCCACACCTATTATTAAGCACTGCCGCTATGAACCAAAATTCAG 718  
 Qy 637 CCCTACTCTCCGCTGTTCCGATATGGGACCTCTGTCGCAAGGCTGGAGGACCTTGA 696  
 Db 719 CCCTACTCTCCGCTGTTCCGATATGGGACCTCTGTCGCAAGGCTGGAGGACCTTGA 778  
 Qy 697 GGACCTGGCTGCTGGCTGGCTCTGTAGGCATCAGAGTTCACTGGGATTGTGACTGGA 756  
 Db 779 GGACCTGGCTGCTGGCTGGCTCTGTAGGCATCAGAGTTCACTGGGATTGTGACTGGA 838  
 Qy 757 CACCGGGACTCTGGCTGCTGGCTCTCACTACTCTTCCAGCTGCAGGAGAGAGCTACAA 816  
 Db 839 CACCGGGACTCTGGCTGCTGGCTCTCACTACTCTTCCAGCTGCAGGAGAGAGCTACAA 898  
 Qy 817 CTTACGACGACCTCACTGCTGGGACCAACCGGTGTGGAGGCCCGCACCTCTCTCAA 876  
 Db 899 CTTACGACGACCTCACTGCTGGGACCAACCGGTGTGGAGGCCCGCACCTCTCTCAA 958  
 Qy 877 GCTCTATGAATCCGCTTTCGACATCTCTGTCACCGGGCAGGAGGAAAGTTGGGCTCAT 936  
 Db 959 GCTCTATGAATCCGCTTTCGACATCTCTGTCACCGGGCAGGAGGAAAGTTGGGCTCAT 1018  
 Qy 937 CCCCACGGCGTCACACTGGGACCGGGCAGCTTGGCTGGGCGTGTCACTTTTCTG 996  
 Db 1019 CCCCACGGCGTCACACTGGGACCGGGCAGCTTGGCTGGGCGTGTCACTTTTCTG 1078  
 Qy 997 TGACCTGCTACTGCTGTATGTGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1056  
 Db 1079 TGACCTGCTACTGCTGTATGTGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1138  
 Qy 1057 GGAGGCCAAGGCCCGGAAAGCAACCGCACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1116  
 Db 1139 GGAGGCCAAGGCCCGGAAAGCAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1198  
 Qy 1117 CCAAGCCCACTGGCCGAGTGCTCAGACGAGCTCAGACCTGACACCCACGCGCACTGC 1176  
 Db 1199 CCAAGCCCACTGGCCGAGTGCTCAGACGAGCTCAGACCTGACACCCACGCGCACTGC 1258  
 Qy 1177 TGTGGAGTCTAGACACAGACACCAAGGATGGCCCTGTCCAAATTTCTGACACCCACTTGC 1236  
 Db 1259 TGTGGAGTCTAGACACAGACACCAAGGATGGCCCTGTCCAAATTTCTGACACCCACTTGC 1318  
 Qy 1237 AACCCATTCCGGAGCTGTAGCCG-TTCCCTGCTGTTGAGA-----GTTGGGG 1286  
 Db 1319 AACCCATTCCGGAGCTGTAGCCGTTTCCCTGCTGTTGAGAAGAGAGAGAGGGGCTGGGC 1378  
 Qy 1287 CTGGGAAGGGCGGGGCCCTGCTGGGATCTCAAGGATGAGGCCCCAGCATGGA 1340  
 Db 1379 AAGGAAGGACCCCTGCTGCCCTGCCGAGCGAAGCAAGGATGAGGCAACAGCAATGA 1432

## RESULT 7

ADP49174

ID ADP49174 standard; DNA; 1452 BP.

AC ADP49174;

XX ADP49174;

XX 26-AUG-2004 (first entry)

DT 26-AUG-2004 (first entry)

XX Rat P2X2 purinergic receptor DNA sequence for odour modulation.  
 DE DE

XX ds; gene; odour sensitivity; P2X purinergic receptor;  
 KW P2Y purinergic receptor; signal transduction pathway;  
 KW Olfactory signalling; micro-array.  
 XX

OS Rattus sp.  
 XX  
 PN W02004047749-A2.

XX 10-JUN-2004.  
 PD 21-NOV-2003; 2003WO-US037389.  
 PF 21-NOV-2002; 2002US-0428140P.  
 PR (UTAH ) UNIV UTAH RES FOUND.  
 PA Lucero M, Hegg C;  
 PI WPI; 2004-460642/43.  
 DR Modulating odor sensitivity in a subject, comprises administering a  
 PT composition comprising an agonist or antagonist of P2X or P2Y purinergic  
 PT receptor to the subject.  
 XX Disclosure; SEQ ID NO 3; 108pp; English.  
 XX The invention relates to a method of modulating (M1) odour sensitivity in  
 CC a subject, by administering a composition which is an agonist or  
 CC antagonist of a P2X or P2Y purinergic receptor. (M1) is useful for  
 CC modulating odour sensitivity in a subject (claimed). The compositions  
 CC used for modulating odour sensitivity in a subject are useful for  
 CC studying the signal transduction pathways related to olfactory signaling.  
 CC The compositions are also useful as reagents in micro-arrays or as  
 CC reagents to probe or analyze existing micro-arrays. This sequence  
 CC corresponds to the rat P2X2 DNA sequence.  
 XX Sequence 1452 BP; 316 A; 415 C; 442 G; 279 T; 0 U; 0 Other;  
 SQ

Query Match 43.6%; Score 1174; DB 12; Length 1452;  
 Best Local Similarity 97.2%; Pred. No. 3.2e-267;  
 Matches 1219; Conservative 0; Mismatches 25; Indels 10; Gaps 2;

Qy 97 GTGGGCTCTCTCCCAAAAAGGCTACCAAGGCGGAGCTGGAAACCCAGTTTCCAT 156  
 Db 179 GTGGGCTCTCTCCCAAAAAGGCTACCAAGGCGGAGCTGGAAACCCAGTTTCCAT 238  
 Qy 157 CATCAACAACTCAAAAGGGGTTCCGTCACATCAAGAGAGCTTGGAAACCGGCTGTG 216  
 Db 239 CATCAACAACTCAAAAGGGGTTCCGTCACATCAAGAGAGCTTGGAAACCGGCTGTG 298  
 Qy 217 GGATGTGGCGACTTCGTGAAGCCACTCAGGGAGAGAACGTGTTCTTGTGTGACAA 276  
 Db 299 GGATGTGGCGACTTCGTGAAGCCACTCAGGGAGAGAACGTGTTCTTGTGTGACAA 358  
 Qy 277 CTTCCTTGTGACGCCACCCCAAGTTCAAGGAGATGCCAGAGACCCGCTCCGCTCCACT 336  
 Db 359 CTTCCTTGTGACGCCACCCCAAGTTCAAGGAGATGCCAGAGACCCGCTCCGCTCCACT 418  
 Qy 337 GGCTAACTGCTGGGTGACAGGAGACTGCCCCGAAAGGGAGGAGGCACACAGCCACGG 396  
 Db 419 GGCTAACTGCTGGGTGACAGGAGACTGCCCCGAAAGGGAGGAGGCACACAGCCACGG 478  
 Qy 397 TGTAAAACAGGCCAGTGTGTGTTCAATGGAGCCCAAGGACCTGTGAGTCTCGAG 456  
 Db 479 TGTAAAACAGGCCAGTGTGTGTTCAATGGAGCCCAAGGACCTGTGAGTCTCGAG 538  
 Qy 457 TTGCTGCCCAAGTGGAGAGTGGCGTTGTGCCCTCAGAGCCCTCTGTGGCCCGCCAGAA 516  
 Db 539 TTGCTGCCCAAGTGGAGAGTGGCGTTGTGCCCTCAGAGCCCTCTGTGGCCCGCCAGAA 598  
 Qy 517 CTTCACTGTTTCATCAAAAACACAGTCACCTTCAGCAAGTTCACTTCTTAAGTCAA 576  
 Db 599 CTTCACTGTTTCATCAAAAACACAGTCACCTTCAGCAAGTTCACTTCTTAAGTCAA 658  
 Qy 577 TGCCTTGGAGACCTGGGACCCCACTATTTTAAGCACTGCGGCTATGAACCAAAATTCAG 636  
 Db 659 TGCCTTGGAGACCTGGGACCCCACTATTTTAAGCACTGCGGCTATGAACCAAAATTCAG 718  
 Qy 637 CCCCTACTGTCCTGTTCCGCAATTTGGGAGCCTGCTGGCCAAAGGCTGGAGGACCTTCGA 696

Db 719 CCCTACTGTCCTGCTGCTCCGATTCGGGACCTCGTGCCCAAGCTGGAGGACCTTCGA 778  
 QY 697 GCACCTGGGTTGCTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTCGACCTGGA 756  
 Db 779 GGACCTGGGTTGCTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTCGACCTGGA 838  
 QY 757 CACCGGGGACTCTGGCTGCTGGCTCTCACTACTCTCTCCAGCTGCAGGAGAGAGCTCAA 816  
 Db 839 CACCGGGGACTCTGGCTGCTGGCTCTCACTACTCTCTCCAGCTGCAGGAGAGAGCTCAA 898  
 QY 817 CTTTACGACAGCACTCACTGTGGGACCAACCGGCTGTGAGGCCCGACCTGCTCAA 876  
 Db 899 CTTTACGACAGCACTCACTGTGGGACCAACCGGCTGTGAGGCCCGACCTGCTCAA 958  
 QY 877 GCTCTATGAATCCGCTTCGACATCCTCGTCAACCGGCGAGGAGAGTTTCGGGCTCAT 936  
 Db 959 GCTCTATGAATCCGCTTCGACATCCTCGTCAACCGGCGAGGAGAGTTTCGGGCTCAT 1018  
 QY 937 CCCACGGCCGTCACACTGGGACCCGGGGAGCTTGGCTGGGGTGGTGCACCTTTTCTG 996  
 Db 1019 CCCCACGGCCGTCACACTGGGACCCGGGGAGCTTGGCTGGGGTGGTGCACCTTTTCTG 1078  
 QY 997 TGACCTGCTACTGCTGTATGTGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1056  
 Db 1079 TGACCTGCTACTGCTGTATGTGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1138  
 QY 1057 GGAGGCCAAGGCCCGAAGCAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTTGCATC 1116  
 Db 1139 GGAGGCCAAGGCCCGAAGCAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTTGCATC 1198  
 QY 1117 CCAACGCCGACTGGCCGAGTGCCTCAGACGGAGCTCAGCACTGCACCCAGGCCACTGC 1176  
 Db 1199 CCAACGCCGACTGGCCGAGTGCCTCAGACGGAGCTCAGCACTGCACCCAGGCCACTGC 1258  
 QY 1177 TGCTGGGAGTCAGACACAGACACAGGATGGCCCTGTCCAAAGTTCTGACACCCACTTGCC 1236  
 Db 1259 TGCTGGGAGTCAGACACAGACACAGGATGGCCCTGTCCAAAGTTCTGACACCCACTTGCC 1318  
 QY 1237 AACCCATTCGGGAGCTGTAGCCG-TTCCCTGTGCTGGTTGAGA-----GTTGGGGG 1286  
 Db 1319 AACCCATTCGGGAGCTGTAGCCGTTTCCCTGTGCTGGTTGAGAAGAGAGGGGCTGGG 1378  
 QY 1287 CTGGAGGGGGGGGCTGCTGGGATCTCAAGGATGAGGCCCGCCAGCATGA 1340  
 Db 1379 AAGGAGGACCCCTGCCCTGCGGAGCAAGCAAGGATGAGGCCAGCATGA 1432

## RESULT 8

ADP49176

ID ADP49176 standard; DNA; 1452 Bp.

AC ADP49176;

XX ADP49176;

DT 26-AUG-2004 (first entry)

DE Human P2X3 purinergic receptor DNA sequence for odour modulation.

XX ds; gene; odour sensitivity; P2X purinergic receptor;

XX P2Y purinergic receptor; signal transduction pathway;

XX olfactory signalling; micro-array.

XX Homo sapiens.

OS Homo sapiens.

FN WO2004047749-A2.

PD 10-JUN-2004.

XX 21-NOV-2003; 2003WO-US037389.

XX 21-NOV-2002; 2002US-0428140P.

XX (UTAH) UNIV UTAH RES FOUND.

XX

XX

XX

XX

XX

XX

PI Lucero M, Hegg C;  
 XX WPI; 2004-460642/43.  
 XX Modulating odor sensitivity in a subject, comprises administering a  
 PT composition comprising an agonist or antagonist of P2X or P2Y purinergic  
 PR receptor to the subject.  
 XX Disclosure; SEQ ID NO 5; 108pp; English.  
 XX  
 CC The invention relates to a method of modulating (M1) odour sensitivity in  
 CC a subject, by administering a composition which is an agonist or  
 CC antagonist of a P2X or P2Y purinergic receptor. (M1) is useful for  
 CC modulating odour sensitivity in a subject (claimed). The compositions  
 CC used for modulating odour sensitivity in a subject are useful for  
 CC studying the signal transduction pathways related to olfactory signaling.  
 CC The compositions are also useful as reagents in micro-arrays or as  
 CC reagents to probe or analyze existing micro-arrays. This sequence  
 CC corresponds to the human P2X3 DNA sequence.  
 XX  
 SQ Sequence 1452 Bp; 316 A; 415 C; 442 G; 279 T; 0 U; 0 Other;  
 Query Match 43.6%; Score 1174; DB 12; Length 1452;  
 Best Local Similarity 97.2%; Pred. No. 3.2e-267;  
 Matches 1219; Conservative 0; Mismatches 25; Indels 10; Gaps 2;  
 QY 97 GTGGGCTCTCTCGCCAAAAGGCTACAGAGCGGACCTGGAAACCCAGTTTCCAT 156  
 Db 179 GTGGGCTCTCTCGCCAAAAGGCTACAGAGCGGACCTGGAAACCCAGTTTCCAT 238  
 QY 157 CATCACCAACTCAAGGGGTTTCGTCACCTCAGATCAAGGAGCTTGAACCGGCTGTG 216  
 Db 239 CATCACCAACTCAAGGGGTTTCGTCACCTCAGATCAAGGAGCTTGAACCGGCTGTG 298  
 QY 217 GGATGTGGCGGACTCTCGTGAAGCCACCTCAGGAGAGAACCTGTTCCTTCTGGTGACCAA 276  
 Db 299 GGATGTGGCGGACTCTCGTGAAGCCACCTCAGGAGAGAACCTGTTCCTTCTGGTGACCAA 358  
 QY 277 CTTCTCTGTGAGCCAGCCCAAGTTTCAGGAGAGATGCCAGAGACCCGTCGTCCTCACT 336  
 Db 359 CTTCTCTGTGAGCCAGCCCAAGTTTCAGGAGAGATGCCAGAGACCCGTCGTCCTCACT 418  
 QY 337 GCCTAACTCTCGGTCGAGGAGCTGCCCGGAGGAGGAGGACACACAGCCAGCG 396  
 Db 419 GGCTAACTCTCGGTCGAGGAGCTGCCCGGAGGAGGAGGAGGACACACAGCCAGCG 478  
 QY 397 TGTAAAAACAGCCAGCTGTGGTGTTCATATGGGACCCACAGGACCTGTGAGATCTGGAG 456  
 Db 479 TGTAAAAACAGCCAGCTGTGGTGTTCATATGGGACCCACAGGACCTGTGAGATCTGGAG 538  
 QY 457 TTGGTGGCCAGTGAGAGTGGCGTTGTGCCCTCGAGGCCCCCTGCTGGCCCCAGGCCCCAGAA 516  
 Db 539 TTGGTGGCCAGTGAGAGTGGCGTTGTGCCCTCGAGGCCCCCTGCTGGCCCCAGGCCCCAGAA 598  
 QY 517 CTTCACTGTTCATCAAAAAACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTCCAA 576  
 Db 599 CTTCACTGTTCATCAAAAAACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTCCAA 658  
 QY 577 TGGCTTGGAGACCTGGGACCCCACTATTTAAGCACTGCGGCTATGAACACCAATTCAG 636  
 Db 659 TGGCTTGGAGACCTGGGACCCCACTATTTAAGCACTGCGGCTATGAACACCAATTCAG 718  
 QY 637 CCCCTACTGTCCCGTGTTCGCGATTTGGGACCTCGTGGCCCAAGGCTGGAGGACCTTCGA 696  
 Db 719 CCCCTACTGTCCCGTGTTCGCGATTTGGGACCTCGTGGCCCAAGGCTGGAGGACCTTCGA 778  
 QY 697 GGACCTGGGTTGCTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTCGACCTGGA 756  
 Db 779 GGACCTGGGTTGCTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTCGACCTGGA 838  
 QY 757 CACCGGGGACTCTGGCTGCTGCTCACTACTCTCTCCAGCTGCAGGAGAGAGCTACAA 816  
 Db 839 CACCGGGGACTCTGGCTGCTGCTCACTACTCTCTCCAGCTGCAGGAGAGAGCTACAA 898

```
QY 817 CTTTCAGGACGACCTCACTGTTGGAGCAACCGGGTGGAGGCCCGCACCTCTCTCAA 876
DB 899 CTTTCAGGACGACCTCACTGTTGGAGCAACCGGGTGGAGGCCCGCACCTCTCTCAA 958
QY 877 GCTCTATGAATCCGCTTCGACATCTCTGTCACCGGGCAGGAGGAAAGTTCCGGCTCAT 936
DB 959 GCTCTATGAATCCGCTTCGACATCTCTGTCACCGGGCAGGAGGAAAGTTCCGGCTCAT 1018
QY 937 CCCACGGCGGTACACTGGGCACCGGGCAGCTTGGCTGGGGGTGTACCTTTTCTG 996
DB 1019 CCCACGGCGGTACACTGGGCACCGGGCAGCTTGGCTGGGGGTGTACCTTTTCTG 1078
QY 997 TGACCTGCTACTCTCTGATGTGATAGAGAGCCCAATTTCTACTGGAGGACAAAGTATGA 1056
DB 1079 TGACCTGCTACTCTCTGATGTGATAGAGAGCCCAATTTCTACTGGAGGACAAAGTATGA 1138
QY 1057 GGAGGCGAAGGCCCCGAAAGCAACCGCCAACTCTGTGTGGAGGAGCTGGCCCTTGCAATC 1116
DB 1139 GGAGGCGAAGGCCCCGAAAGCAACCGCCAACTCTGTGTGGAGGAGCTGGCCCTTGCAATC 1198
QY 1117 CCAAGCCGACTGGCGAGTGCTCTCAGCGGAGCTCAGCACCTGCACCCACGCGCACTGC 1176
DB 1199 CCAAGCCGACTGGCGAGTGCTCTCAGCGGAGCTCAGCACCTGCACCCACGCGCACTGC 1258
QY 1177 TGTCTGGAGTACAGACACAGACACAGGATGGCCCTGTCCAAGTTCTGACACCCACTTGC 1236
DB 1259 TGTCTGGAGTACAGACACAGACACAGGATGGCCCTGTCCAAGTTCTGACACCCACTTGC 1318
QY 1237 AACCCATTCGGGAGCTGTAGCGG-TTCCCTGCTGTTGAGA-----GTTGGGG 1286
DB 1319 AACCCATTCGGGAGCTGTAGCGGTTCCCTGCTGTTGAGAAGAGAGAGGGGCTGGGC 1378
QY 1287 CTGGGAAGCGGGGGCCCTGCTCTGGGATCTCAAGGATGAGGCCCCAGCATGGA 1340
DB 1379 AAGGAAGGACCCCTGCTGCTGCCGAGCGAAAGCAAGATGAGGCAACAGCAATGA 1432

RESULT 9
AAD04978
ID AAD04978 standard; cDNA; 1360 BP.
AC AAD04978;
XX
DT 17-JUL-2001 (first entry)
DE Human putative receptor P2X6 cDNA.
KW Human; purinergic receptor; P2X6; antidepressant; vulnerable; hypotensive;
KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;
KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;
KW neuromuscular disease; reproductive system disorder; hypertension;
KW peripheral vascular disease; immune system disorder; chronic bronchitis;
KW irritable bowel disorder; premature ejaculation; asthma; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT misc_feature 1..14
FT /tag= a
FT /note= "5' flanking sequence including BamHI restriction
FT site"
FT CDS 16..1341
FT /tag= b
FT /product= "Human P2X6 receptor protein"
FT 1344..1360
FT /tag= c
FT /note= "3' flanking sequence containing Not I restriction
FT site"
XX
XX US6214581-B1.
XX
XX 10-APR-2001.
XX
XX
```

```
XX 13-NOV-1998; 98US-001911136.
XX
XX 16-JAN-1998; 98US-0008185.
PR 16-JAN-1998; 98US-0008526.
PR 16-JAN-1998; 98US-0071298P.
PR 16-JAN-1998; 98US-0071669P.
XX
PA (ABBO ) ABBOTT LAB.
XX
XX Lynch KJ, Burgard EC, Van Biesen T;
XX
XX WPI; 2001-315459/33.
DR P-PSDB; AAE01141.
XX
XX Novel isolated polynucleotide encoding human purinergic P2X3 receptor
XX polypeptide useful for identifying potentially therapeutic compounds that
XX modulate or otherwise interact with P2X3 containing receptors.
XX
XX Example 9; Fig 8; 53pp; English.
XX
XX The present sequence is a cDNA encoding human P2X6 purinergic receptor
XX protein. P2X receptors are ligand-gated ion channels while P2Y receptors
XX operate generally through a G-protein coupled system. P2X purinoreceptor
XX drugs are potential therapeutic agents in several disorders including
XX central nervous system or peripheral nervous system conditions, e.g.;
XX epilepsy, pain, depression, neurodegenerative disorders, disorders of the
XX skeletal muscle such as neuromuscular diseases, disorders of the
XX reproductive system, asthma, peripheral vascular disease, hypertension,
XX immune system disorders, irritable bowel disorder, premature ejaculation,
XX cystic fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the
XX activity of extracellular nucleotide triphosphates to regulate chloride
XX secretion in human airway epithelia
XX
XX Sequence 1360 BP; 284 A; 403 C; 412 G; 261 T; 0 U; 0 Other;
XX
XX Query Match 43.2%; Score 1162.4; DB 4; Length 1360;
XX Best Local Similarity 99.9%; Pred. No. 1.7e-264;
XX Matches 1163; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 97 GTGGGCTCTCTCCCAAAAAGGCTACAGGACGGAGCTGGAACCCAGTTTCCAT 156
DB 180 GTGGGCTCTCTCCCAAAAAGGCTACAGGACGGAGCTGGAACCCAGTTTCCAT 239
QY 157 CATCACCAACTCAAGGGTTTCCGTCACATCAGATCAAGAGCTTGGAAACCGGCTGT 216
DB 240 CATCACCAACTCAAGGGTTTCCGTCACATCAGATCAAGAGCTTGGAAACCGGCTGT 299
QY 217 GGATGTGGCCGACTTGTGAAGCCACTCAGGGAGAGAACCTGTCTTCTTGTGACCAA 276
DB 300 GGATGTGGCCGACTTGTGAAGCCACTCAGGGAGAGAACCTGTCTTCTTGTGACCAA 359
QY 277 CTTCTTGTGTGAGCGGACCCAAAGTTTCAGGGCAGATGCCAGAGCACCCGTCCTCCACT 336
DB 360 CTTCTTGTGTGAGCGGACCCAAAGTTTCAGGGCAGATGCCAGAGCACCCGTCCTCCACT 419
QY 337 GGCTAACTGTGGTTCGACGAGGACTGCCCGAAGGGAGGAGGAGGCACACAGCACGG 396
DB 420 GGCTAACTGTGGTTCGACGAGGACTGCCCGAAGGGAGGAGGAGGCACACAGCACGG 479
QY 337 TGTAAAAACAGCCAGTGTGTGGTTTCAATGGACCCACAGGACCTGTGTAGATCTGAG 456
DB 480 TGTAAAAACAGCCAGTGTGTGGTTTCAATGGACCCACAGGACCTGTGTAGATCTGAG 539
QY 457 TTGGTGGCCAGTGGAGTGGCGTTGTGCCCTCGAGGCCCTGTGGCCAGGCCAGAA 516
DB 540 TTGGTGGCCAGTGGAGTGGCGTTGTGCCCTCGAGGCCCTGTGGCCAGGCCAGAA 599
QY 517 CTTCACTGTTCATCAAAAAACAGATCACCTTCAGCAAGTTCAACTTCTCTAAGTCAA 576
DB 600 CTTCACTGTTCATCAAAAAACAGATCACCTTCAGCAAGTTCAACTTCTCTAAGTCAA 659
QY 577 TGCCTTGGAGACCTGGGACCCCACTATTTTAAGCACTGCCGCTATGNACCAATTCAG 636
```

```

Db 660 TGCCCTGGAGACCTGGAGACCCACCTATTTTAAGCACTCCCGCTATGAACCAATTCAG 719
QY 637 CCCTACTGTCCCGTGTTCGGCATTTGGGACCTCTGTGGCAAGGCTGGAGGACCTTGA 696
Db 720 CCCTACTGTCCCGTGTTCGGCATTTGGGACCTCTGTGGCAAGGCTGGAGGACCTTGA 779
QY 697 GACCTGGCGTGTCTGGGTGGCTCTGTAGGATCAGAGTTCTCTGGGATTTGACCTGA 756
Db 780 GGACCTGGCGTGTCTGGGTGGCTCTGTAGGATCAGAGTTCTCTGGGATTTGACCTGA 839
QY 757 CACCGGGGACTCTGGTGTGGCTCAGTACTCTCTTCAAGTGCAGGAGAGAGTCAAA 816
Db 840 CACCGGGGACTCTGGTGTGGCTCAGTACTCTCTTCAAGTGCAGGAGAGAGTCAAA 899
QY 817 CTTGAGACAGGCACTCACTGTGGAGCAACCGGCTGTGGAGCCCGACCTCTCAA 876
Db 900 CTTGAGACAGGCACTCACTGTGGAGCAACCGGCTGTGGAGCCCGACCTCTCAA 959
QY 877 GCTCTATGAATCCGCTTCGACATCTCTGTACCCGGGAGGAGGAGGAGTTCGGCTCAT 936
Db 960 GCTCTATGAATCCGCTTCGACATCTCTGTACCCGGGAGGAGGAGGAGTTCGGCTCAT 1019
QY 937 CCCCAGGCGCTCACACTGGGACCCGGGACGCTTTGGCTGGCGTGTCACTTTTCTG 996
Db 1020 CCCCAGGCGCTCACACTGGGACCCGGGACGCTTTGGCTGGCGTGTCACTTTTCTG 1079
QY 997 TGACCTGTACTCTGTATGTGATAGAGAGCCATTTCTACTGGAGACAAATATGA 1056
Db 1080 TGACCTGTACTCTGTATGTGATAGAGAGCCATTTCTACTGGAGACAAATATGA 1139
QY 1057 GGAGGCCAAGGCCCCGAAAGCAACCGCCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1116
Db 1140 GGAGGCCAAGGCCCCGAAAGCAACCGCCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1199
QY 1117 CCAAGCCGACTGGCGGAGTGCCTCAGACGGAGTCAAGCACTGCACCCAGCGGCACTGC 1176
Db 1200 CCAAGCCGACTGGCGGAGTGCCTCAGACGGAGTCAAGCACTGCACCCAGCGGCACTGC 1259
QY 1177 TGCTGGAGTCAAGACACAGACACAGGAGTGGCCCTGTCCAAAGTTCTGACACCCACTTGC 1236
Db 1260 TGCTGGAGTCAAGACACAGACACAGGAGTGGCCCTGTCCAAAGTTCTGACACCCACTTGC 1319
QY 1237 AACCATTCGGGAGCTGTAGCC 1260
Db 1320 AACCATTCGGGAGCTGTAGCC 1343

```

## RESULT 10

```

AAV61832
ID AAV61832 standard; cDNA; 1293 BP.
AC AAV61832;
DT 28-APR-1999 (first entry)
XX

```

DE Coding sequence for human p53 regulated protein, P2XM.

XX Human; p53 regulated protein; P2XM; ATP receptor; RP-2 protein;  
 KW thymocyte; apoptosis; tumour-suppressor gene; rhabdoid sarcoma;  
 KW anticancer drug; ss.

OS Homo sapiens.

XX WO9842835-A1.

PN 01-OCT-1998.

PD 18-MAR-1998; 95WO-JP001146.

PF 26-MAR-1997; 97JP-00093044.

PR (SAKA ) OTSUKA PHARM CO LTD.

```

XX Tokino T, Nakamura Y;
XX WPI; 1998-532006/45.
XX P-PSDB; AAW76434.
XX
XX Human gene P2XM whose transcription is induced by p53 - useful, e.g. for
XX diagnostic purposes and in development of new anticancer drugs.
XX Claim 2; Page 22-23; 43pp; Japanese.
XX
XX This sequence represents the coding sequence for the P2XM protein of the
XX invention. The protein is significantly homologous to: (i) the P2X family
XX of ATP receptors, and (ii) RP-2 protein which is expressed in thymocytes
XX during apoptosis. Transcription of the genes is specifically regulated by
XX the tumour-suppressor gene p53. The P2XM gene is specifically expressed
XX in skeletal muscle and has been localised to chromosome 22q11, an area
XX where mutation and sequence losses frequently occur in rhabdoid sarcomas.
XX The genes may be used for diagnostic purposes (e.g. by detecting changes
XX occurring in the gene in sarcomas), using probes and primers containing
XX or derived from all or part of the genes. The genes may further be used
XX in the development of new anticancer drugs

```

XX SQ Sequence 1293 BP; 271 A; 379 C; 389 G; 254 T; 0 U; 0 Other;

Query Match 42.9%; Score 1155.8; DB 2; Length 1293;

Best Local Similarity 99.8%; Pred. No. 6.1e-263;

Matches 1157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 97 GTGGGCTCTCTCCGCAAAAAGGCTTACCAGGAGCGGACCTGGAAACCCAGTTTCCAT 156
Db 135 GTGGGCGCTCTCCGCAAAAAGGCTTACCAGGAGCGGACCTGGAAACCCAGTTTCCAT 194
QY 157 CATCACCAAACTCAAAGGGTTTCGTCACCTCAGATCAAGGAGCTTGGAAACCGGCTGG 216
Db 195 CATCACCAAACTCAAAGGGTTTCGTCACCTCAGATCAAGGAGCTTGGAAACCGGCTGG 254
QY 217 GATGTGGCGCACTTCGTGAAGCCACCTCAGGAGAGAACGTTCTTCTTGGTGACCAA 276
Db 255 GGATGTGGCGCACTTCGTGAAGCCACCTCAGGAGAGAACGTTCTTCTTGGTGACCAA 314
QY 277 CTTCTCTGTGACGCCAGCCAAAGTTTCCAGGAGAGTCCAGGAGACCCGCTCGTCCCAT 336
Db 315 CTTCTCTGTGACGCCAGCCAAAGTTTCCAGGAGAGTCCAGGAGACCCGCTCGTCCCAT 374
QY 337 GGCTAACTGTGGTTCGACGAGGACTGCCCCGAGGGGAGGAGGCACACAGCAGCACGG 396
Db 375 GGCTAACTGTGGTTCGACGAGGACTGCCCCGAGGGGAGGAGGCACACAGCAGCACGG 434
QY 397 TGTAATAAAGGCGAGTGTGTGGTGTTCATGGGACCCACAGGACTGTGAGATCTGGAG 456
Db 435 TGTAATAAAGGCGAGTGTGTGGTGTTCATGGGACCCACAGGACTGTGAGATCTGGAG 494
QY 457 TTGCTGCCAGTGGAGTGGGTGTGTGCCCTCAGGAGCCCTGCTGGCCAGCCGACAGAA 516
Db 495 TTGCTGCCAGTGGAGTGGGTGTGTGCCCTCAGGAGCCCTGCTGGCCAGCCGACAGAA 554
QY 517 CTTACACTGTTTCATCAAAAACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTCCAA 576
Db 555 CTTACACTGTTTCATCAAAAACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTCCAA 614
QY 577 TGCCTTGGAGACTGGGACCCCACTATTTTAAGCACTGCCGCTATGACCACTTTCAG 636
Db 615 TGCCTTGGAGACTGGGACCCCACTATTTTAAGCACTGCCGCTATGACCACTTTCAG 674
QY 637 CCCCCTACTGTCCCGTGTTCGCAATTTGGGACCTCTGTGGCCAAAGGCTGGAGGACCTTCA 696
Db 675 CCCCCTACTGTCCCGTGTTCGCAATTTGGGACCTCTGTGGCCAAAGGCTGGAGGACCTTCA 734
QY 697 GGACCTGGCGTGTCTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTTGACCTGGA 756
Db 735 GGACCTGGCGTGTCTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTTGACCTGGA 794

```

QY	757	CACCGGGAGCTCTGGCTGCTGGCTCCTACTCTCTTCCAGCTGCAGGAGAGAGCTACAA	816
Db	795	CACCGGGAGCTCTGGCTGCTGGCTCCTACTCTCTTCCAGCTGCAGGAGAGAGCTACAA	854
QY	817	CTTCAGGACAGCCACTCACTGGTGGAGCAACCGGGTGTGAGCGCCGACCCCTGCTCAA	876
Db	855	CTTCAGGACAGCCACTCACTGGTGGAGCAACCGGGTGTGAGCGCCGACCCCTGCTCAA	914
QY	877	GCTCTATGGAATCCGCTTCGACATCTCTGTCACCGGGCAGCAGGGAAGTTTCGGGCTCAT	936
Db	915	GCTCTATGGAATCCGCTTCGACATCTCTGTCACCGGGCAGCAGGGAAGTTTCGGGCTCAT	974
QY	937	CCCACGGCGCTCAGACTGGGACCGGGGAGCTTGGCTGGCGGTGTCACCTTTTCTG	996
Db	975	CCCACGGCGCTCAGACTGGGACCGGGGAGCTTGGCTGGCGGTGTCACCTTTTCTG	1034
QY	997	TGACCTGCTACTGCTGTATGTGATAGAGAGAGCCCATTTCTACTGGAGGAGCAAAAGTATGA	1056
Db	1035	TGACCTGCTACTGCTGTATGTGATAGAGAGAGCCCATTTCTACTGGAGGAGCAAAAGTATGA	1094
QY	1057	GGAGGCCAAGGCCCGCAAGCAACCGCCAACTCTGTGTGGAGGAGCTGGCGCTTGCATC	1116
Db	1095	GGAGGCCAAGGCCCGCAAGCAACCGCCAACTCTGTGTGGAGGAGCTGGCGCTTGCATC	1154
QY	1117	CCAAGCCGACTGGCGGAGTGCTCTCAGACGAGGCTCAGCCTGCACCCAGCGGCACCTGC	1176
Db	1155	CCAAGCCGACTGGCGGAGTGCTCTCAGACGAGGCTCAGCCTGCACCCAGCGGCACCTGC	1214
QY	1177	TGCTGGAGTCAAGACACAGACACACGAGGATGGCGCTGTCCAAAGTTCTGACACCCACTTGC	1236
Db	1215	TGCTGGAGTCAAGACACAGACACACGAGGATGGCGCTGTCCAAAGTTCTGACACCCACTTGC	1274
QY	1237	AACCCATTCGGGAGCGCTG	1255
Db	1275	AACCCATTCGGGAGCGCTG	1293
RESULT 11			
ADP28208			
ID	ADP28208 standard; DNA; 900 BP.		
XX			
AC	ADP28208;		
XX			
DT	12-AUG-2004 (first entry)		
XX			
DE	Human secreted protein encoding sequence SEQ ID #206.		
XX			
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;		
KW	cancer; inflammatory; immune; ds; human secreted protein.		
OS	Homo sapiens.		
XX			
PN	WC004035732-A2.		
XX			
PD	29-APR-2004.		
XX			
PF	28-AUG-2003; 2003WO-US026780.		
XX			
PR	29-AUG-2002; 2002US-0406576P.		
PR	29-AUG-2002; 2002US-0406579P.		
PR	29-AUG-2002; 2002US-0406585P.		
PR	29-AUG-2002; 2002US-0406588P.		
PR	29-AUG-2002; 2002US-0406608P.		
PR	29-AUG-2002; 2002US-0406611P.		
PR	29-AUG-2002; 2002US-0406612P.		
PR	29-AUG-2002; 2002US-0406616P.		
PR	29-AUG-2002; 2002US-0406640P.		
PR	29-AUG-2002; 2002US-0406642P.		
PR	29-AUG-2002; 2002US-0406646P.		
PR	29-AUG-2002; 2002US-0406653P.		
PR	29-AUG-2002; 2002US-0406655P.		
PR	29-AUG-2002; 2002US-0406666P.		
PR	17-SEP-2002; 2002US-0410946P.		

PR	17-SEP-2002; 2002US-0410947P.	
PR	17-SEP-2002; 2002US-0410948P.	
PR	17-SEP-2002; 2002US-0410949P.	
PR	17-SEP-2002; 2002US-0410953P.	
PR	17-SEP-2002; 2002US-0410957P.	
PR	17-SEP-2002; 2002US-0410958P.	
PR	17-SEP-2002; 2002US-0410959P.	
PR	17-SEP-2002; 2002US-0410960P.	
PR	17-SEP-2002; 2002US-0410961P.	
PR	17-SEP-2002; 2002US-0410962P.	
PR	17-SEP-2002; 2002US-0411019P.	
PR	17-SEP-2002; 2002US-0411022P.	
PR	17-SEP-2002; 2002US-0411023P.	
PR	17-SEP-2002; 2002US-0411024P.	
PR	17-SEP-2002; 2002US-0411032P.	
PR	17-SEP-2002; 2002US-0411035P.	
PR	17-SEP-2002; 2002US-0411037P.	
PR	17-SEP-2002; 2002US-0411041P.	
PR	17-SEP-2002; 2002US-0411045P.	
PR	17-SEP-2002; 2002US-0411046P.	
PR	17-SEP-2002; 2002US-0411048P.	
PR	17-SEP-2002; 2002US-0411052P.	
PR	17-SEP-2002; 2002US-0411055P.	
PR	17-SEP-2002; 2002US-0411073P.	
PR	17-SEP-2002; 2002US-0411082P.	
PR	17-SEP-2002; 2002US-0411101P.	
PR	18-APR-2003; 2003US-0463700P.	
PR	18-APR-2003; 2003US-0463708P.	
PR	18-APR-2003; 2003US-0463716P.	
PR	18-APR-2003; 2003US-0463732P.	
PR	02-MAY-2003; 2003US-0467199P.	
PR	02-MAY-2003; 2003US-0467201P.	
PR	02-MAY-2003; 2003US-0467203P.	
PR	02-MAY-2003; 2003US-0467230P.	
PR	19-MAY-2003; 2003US-0471306P.	
PR	19-MAY-2003; 2003US-0471356P.	
PR	22-MAY-2003; 2003US-0472420P.	
PR	22-MAY-2003; 2003US-0472430P.	
PR	09-JUN-2003; 2003US-0476609P.	
PR	09-JUN-2003; 2003US-0476641P.	
PR	08-JUL-2003; 2003US-0485218P.	
PR	08-JUL-2003; 2003US-0485223P.	
PR	08-JUL-2003; 2003US-0485224P.	
PR	08-JUL-2003; 2003US-0485325P.	
PR	14-JUL-2003; 2003US-0486446P.	
PR	14-JUL-2003; 2003US-0486480P.	
PR	15-JUL-2003; 2003US-0486891P.	
PR	15-JUL-2003; 2003US-0486960P.	
PR	08-AUG-2003; 2003US-0493341P.	
PR	08-AUG-2003; 2003US-0493370P.	
PR	08-AUG-2003; 2003US-0493573P.	
PR	08-AUG-2003; 2003US-0493577P.	
XX	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
PA	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	
PI	Halenbeck RF, Huang MM, Kothakota S, Linnemann T;	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX	WPI; 2004-348438/32.	
DR		
XX		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX	Claim 1; SEQ ID NO 206; 428pp; English.	
PS		
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	



CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein encoding sequence. The  
 CC present sequence is available on WIPWEB and is not in the specification.  
 XX  
 SQ Sequence 900 BP; 182 A; 251 C; 288 G; 179 T; 0 U; 0 Other;

Query Match 25.2%; Score 678.4; DB 12; Length 900;  
 Best Local Similarity 89.4%; Pred. No. 4.1e-150;  
 Matches 792; Conservative 0; Mismatches 1; Indels 93; Gaps 2;

QY 97 GTGGGCTCTCTCCGCAAAAGGCTACAGAGCGGACCTGGACCCACCTTTTCCAT 156  
 DB 108 GTGGGCTCTCTCCGCAAAAGGCTACAGAGCGGACCTGGACCCACCTTTTCCAT 167

QY 157 CATCACAACTCAAAAGGGTTTCCTACTCAGATCAAGAGCTTGGAAACCGCTGTG 216  
 DB 168 CATCACAACTCAAAAGGGTTTCCTACTCAGATCAAGAGCTTGGAAACCGCTGTG 227

QY 217 GGATGGCGGACTTCGTGAAGCCCTCAGGAGAGACGTGTTCTTTGGTGACCA 276  
 DB 228 GGATGGCGGACTTCGTGAAGCCCTCAGGAGAGACGTGTTCTTTGGTGACCA 287

QY 277 CTTCTTTGACGCCAGCCCAAGTTTCAGGGCAGATGCCAGAGACCCGTCGTCCTCACT 336  
 DB 288 CTTCTTTGACGCCAGCCCAAGTTTCAGGGCAGATGCCAGAGACCCGTCGTCCTCACT 347

QY 337 GGCTAACTGCTGGTTCGACGAGACTGCCCGAAGGGAGGAGGACACACAGCCAGG 396  
 DB 348 GGCTAACTGCTGGTTCGACGAGACTGCCCGAAGGGAGGAGGACACACAGCCAGG 407

QY 397 TGTAAAAACAGCCGAGTGTGCTGTTCAATGGAGCCACAGAGACCTGTGAGATCTGGAG 456  
 DB 408 TGTAAAAACAGCCGAGTGTGCTGTTCAATGGAGCCACAGAGACCTGTGAGATCTGGAG 467

QY 457 TTGGTCCGAGTGAGAGTGCGTTGTGCCCTCGAGGCCCTGTGTCGCCAGGCCAGAA 516  
 DB 468 TTGGTCCGAGTGAGAGTGCGTTGTGCCCTCGAGGCCCTGTGTCGCCAGGCCAGAA 524

QY 517 CTTACACTGTTTCATCAAAAAACAGTCACTTCCAGCAAGTTCACATCTCTAGTCAA 576  
 DB 525 CTTACACTGTTTCATCAAAAAACAGTCACTTCCAGCAAGTTCACATCTCTAGTCAA 566

QY 577 TGCCTTGAGACCTGGGACCCCACTATTTAAGCACTGCGCTATGAACCAATTCAG 636  
 DB 567 ----- 566

QY 637 CCCCTACTGTCCTGTTTCGATTTGGGACCTGTGCGCAAGCTCGAGGGACCTTCGA 696  
 DB 567 -----CGTGTTCGCAATGGGACCTGTGCGCAAGCTCGAGGGACCTTCGA 614

QY 697 GGACTGCGTGTGCTGGTGGTCTCTAGGCATCAGAGTTCACTGGGATTTGACCTGGA 756  
 DB 615 GGACTGCGTGTGCTGGTGGTCTCTAGGCATCAGAGTTCACTGGGATTTGACCTGGA 674

QY 757 CACCGGGACTCTGGCTGGCTCTACTCTCTCAGCTGAGAGAGAGTACAA 816  
 DB 675 CACCGGGACTCTGGCTGGCTCTACTCTCTCAGCTGAGAGAGAGTACAA 734

QY 817 CTTCAGGACAGCCACTCACTGTTGGGAGCAACCGGTTGTGAGCCCGCACCTGCTCAA 876  
 DB 735 CTTCAGGACAGCCACTCACTGTTGGGAGCAACCGGTTGTGAGCCCGCACCTGCTCAA 794

QY 877 GCTCTATGGAATCCGCTTCGACATCTCTGTCACCGGACAGGAGGAGTTCCGGGCTCAT 936  
 DB 795 GCTCTATGGAATCCGCTTCGACATCTCTGTCACCGGACAGGAGGAGTTCCGGGCTCAT 854

QY 937 CCCCACGGCGCTCAGCTGGGACCGGGGAGGCTTGGCTGGGCGTG 982  
 DB 855 CCCCACGGCGCTCAGCTGGGACCGGGGAGGCTTGGCTGGGCGTG 900

ID  
 XX AA118520 standard; DNA; 569 BP.  
 AC AA118520;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Probe #8453 for gene expression analysis in human cervical cell sample.  
 CC Probe: human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer; ss.  
 XX  
 OS Homo sapiens.  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234887P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human cervical epithelial cells.  
 XX  
 XX Claim 25; SEQ ID NO 8453; 487pp; English.  
 XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENPs). The present sequence is one such probe. The SENPs are derived  
 CC from human HeLa cells. The SENPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging of  
 CC diseases of the cervix, notably cervical cancer. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 569 BP; 131 A; 167 C; 167 G; 104 T; 0 U; 0 Other;

Query Match 21.1%; Score 567.4; DB 4; Length 569;  
 Best Local Similarity 99.8%; Pred. No. 6.2e-124;  
 Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1497 ACCTGTATTGAGGGCTCCGACTGCATGTGCAGGGGCTCTGCTGGTCTGGGCTTGA 1556  
 DB 569 ACCTGTATTGAGGGCTCCGACTGCATGTGCAGGGGCTCTGCTGGTCTGGGCTTGA 510

QY 1557 GGTCTCTCTCCAGTGTCTTCTCCCAAGTGTCTTAGCAGAGGTATGCTTACAGCTGTC 1616  
 DB 509 GGTCTCTCTCCAGTGTCTTCTCCCAAGTGTCTTAGCAGAGGTATGCTTACAGCTGTC 450

QY 1617 AGCAGAGACCTCTGCTGGTGGTCTGCGCCCTCTCCCATCTGACCCCATCAT 1676  
 DB 449 AGCAGAGACCTCTGCTGGTGGTCTGCGCCCTCTCCCATCTGACCCCATCAT 390

QY 1677 AGGTAGAGACCCCACTCTCCCATCTGCTTACATGGGGCTGTGAGCTGGAGCCAAAAG 1736  
 DB 389 AGGTAGAGACCCCACTCTCCCATCTGCTTACATGGGGCTGTGAGCTGGAGCCAAAAG 330

QY 1737 GCAGGGCAGAAAGAGAGTGTATGGGGAGGGGATGTTTTCAGCTTCTCTGGTCTGTA 1796



Db 329 GCAAGTAGAAGAGGAGTATGGGGAGGGGATTGTTTCAGCTTCTCTGGTGTGTGA 270  
Qy 1797 TGCCCCAGGAGAGTCTTAATCTAGGGAATGGGTGAGTAGCGAGATAATCCACTCCCT 1856  
Db 269 TGCCCCAGGAGAGTCTTAATCTAGGGAATGGGTGAGTAGCGAGATAATCCACTCCCT 210  
Qy 1857 ATCCCCAGGAGGAGGAGCATGCTTTGGGCCACACCTGCTTAGTTTATGAGGAC 1916  
Db 209 ATCCCCAGGAGGAGGAGCATGCTTTGGGCCACACCTGCTTAGTTTATGAGGAC 150  
Qy 1917 CGGCTGCTTTCCAGTGGTAGCCCTTTTGCATGAGGAGTCTGGAGAGAGACAGAGGCG 1976  
Db 149 CGGCTGCTTTCCAGTGGTAGCCCTTTTGCATGAGGAGTCTGGAGAGAGACAGAGGCG 90  
Qy 1977 GAGGGCTAAGTGTGTGATCATGGGTCTTCAGACCTTCTATATCCCTCTCGGTAAC 2036  
Db 89 GAGGGCTAAGTGTGTGATCATGGGTCTTCAGACCTTCTATATCCCTCTCGGTAAC 30  
Qy 2037 CCCCCAGGCCAACCCCTTGGAAATCTTCC 2065  
Db 29 CCCCCAGGCCAACCCCTTGGAAATCTTCC 1

## RESULT 13

ABA63528/c

ID ABA63528 standard; DNA; 569 BP.

XX ABA63528;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #11833.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234587P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/S2.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human foetal liver.

XX Claim 1; SEQ ID NO 11833; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring

XX human gene expression in a sample derived from human foetal liver. The

XX single exon nucleic acid probes may be used for predicting, measuring and

XX displaying gene expression in samples derived from human foetal liver. The

XX present sequence is a single exon nucleic acid probe of the invention.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 569 BP; 131 A; 167 C; 167 G; 104 T; 0 U; 0 Other;

Query Match 21.1%; Score 567.4; DB 4; Length 569;  
Best Local Similarity 99.8%; Pred. No. 6.2e-124;  
Matches 569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1497 ACTGTATTGACGGGCTCCGACTGTGATGTGGAGGGGCTCTGCTGCTGCTGGGCTGGA 1556  
Db 569 ACTGTATTGACGGGCTCCGACTGTGATGTGGAGGGGCTCTGCTGCTGCTGGGCTGGA 510  
Qy 1557 GGTCCTCTCCCAAGTCTGCTCCAGTGTTCCTAGCAGAGATGCTTACCAGTGTG 1616  
Db 509 GGTCCTCTCTCCCAAGTCTGCTCCAGTGTTCCTAGCAGAGATGCTTACCAGTGTG 450  
Qy 1617 AGCAGAGACCTCTGCTGCTGGGTCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1576  
Db 449 AGCAGAGACCTCTGCTGCTGGGTCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 390  
Qy 1677 AGGTAGAGACCCACCTCCATCGGTCTTACATGGGGCTGTGCTGCTGCTGCTGCTGCT 1736  
Db 389 AGGTAGAGACCCACCTCCATCGGTCTTACATGGGGCTGTGCTGCTGCTGCTGCTGCT 330  
Qy 1737 GCAAGGCAGAAAGAGAGAGTGTATGGGGAGGGGGAATGTTTTCAGCTTCTCTGGTGTG 1796  
Db 329 GCAAGGTAGAAAGAGAGAGTGTATGGGGAGGGGGAATGTTTTCAGCTTCTCTGGTGTG 270  
Qy 1797 TGCCCCAGGAGAGTCTTAATCTAGGGAATGGGTGAGTAGCGAGATAATCCACTCCCT 1856  
Db 269 TGCCCCAGGAGAGTCTTAATCTAGGGAATGGGTGAGTAGCGAGATAATCCACTCCCT 210  
Qy 1857 ATCCCCAGGAGGAGGAGCATGCTTCTGGGCCACACCTGCTTGTAGTTATGAGGAC 1916  
Db 209 ATCCCCAGGAGGAGGAGCATGCTTCTGGGCCACACCTGCTTGTAGTTATGAGGAC 150  
Qy 1917 CGGCTGCTTTCCAGTGGTAGCCCTTTTGCATGAGGAGTCTGGAGAGAGACAGAGGCG 1976  
Db 149 CGGCTGCTTTCCAGTGGTAGCCCTTTTGCATGAGGAGTCTGGAGAGAGACAGAGGCG 90  
Qy 1977 GCAGGGCTAAGTGTGTGATCATGGGTCTTCAGACCTTCTATATCCCTCTCGGTAAC 2036  
Db 89 GCAGGGCTAAGTGTGTGATCATGGGTCTTCAGACCTTCTATATCCCTCTCGGTAAC 30  
Qy 2037 CCCCCAGGCCAACCCCTTGGAAATCTTCC 2065  
Db 29 CCCCCAGGCCAACCCCTTGGAAATCTTCC 1

## RESULT 14

AAI43639/c

ID AAI43639 standard; DNA; 569 BP.

XX AAI43639;

XX 17-OCT-2001 (first entry)

DT Probe #12325 used to measure gene expression in human placenta sample.  
DE Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder; ss.  
KW Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234587P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.



Db 269 TGCCCCAGGAGTCCCTAATCTAGGAATGGGTGGAGTAGGAGATATATCCACCTCCCT 210  
QY 1857 ATCCCCCAGGCAAGGGCGGAGCATGTCTTTGGGCCCCACACCTGTCTTAGTTTATGAGGAC 1916  
Db 209 ATCCCCCAGGCAAGGGCGGAGCATGTCTTTGGGCCCCACACCTGTCTTAGTTTATGAGGAC 150  
QY 1917 CGGCTGCTTCCAGTCGTAGCCCTTTTGCCATGGAGGTCTGGAGAGAGAGCAGAGGGCG 1976  
Db 149 CGGCTGCTTCCAGTCGTAGCCCTTTTGCCATGGAGGTCTGGAGAGAGAGCAGAGGGCG 90  
QY 1977 GCAGGGCTAAGTTGGTGATCATTTGGGTCTTCAGGACCTTCTATATCCCTCCGTAAAC 2036  
Db 89 GCAGGGCTAAGTTGGTGATCATTTGGGTCTTCAGGACCTTCTATATCCCTCCGTAAAC 30  
QY 2037 CCCCCAGGCCAACCCCTTGGAACTCTTCC 2065  
Db 29 CCCCCAGGCCAACCCCTTGGAACTCTTCC 1

Search completed: November 21, 2004, 18:28:41  
Job time : 1297.69 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 12:30:45 ; Search time 11523.5 Seconds  
(without alignments)  
11051.408 Million cell updates/sec

Title: US-09-820-095B-1  
Perfect score: 2693  
Sequence: 1 ttgtgtactcatgtgccgc.....aaaaaaaaaaaaaaaaaaaaa 2693

Scoring table: IDENTITY NUC  
Gap 10.0, Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053459

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2585.2	96.0	2738	BC033488	BC033488 Homo sapi
2	2558.8	95.0	3554	CQ730055	CQ730055 Sequence
3	2517.2	93.5	3552	AB002058	AB002058 Homo sapi
4	2382.8	88.5	2578	BC064805	BC064805 Homo sapi
5	1592.6	59.1	147086	AC002472	AC002472 Homo sapi
6	1592.6	59.1	162470	AC007664	AC007664 Homo sapi
7	1552.6	57.7	28984	AB002059	AB002059 Homo sapi
8	1515.4	56.3	1697	AR161588	AR161588 Sequence
9	1515.4	56.3	1697	AR161588	AR161588 Human mRNA
10	1336.8	49.6	1517	CR456535	CR456535 Homo sapi
11	1204.4	44.7	2034	BC063553	BC063553 Homo sapi
12	1174	43.6	1452	AF065385	AF065385 Homo sapi
13	1155.8	42.9	1293	AR161587	AR161587 Sequence
14	1155.8	42.9	1293	AR161587	AR161587 Human mRNA
15	777	28.9	1484	AB010883	AB010883 Mus muscu
16	731.4	27.2	2243	X92070	X92070 R.norvegicu
17	729.2	27.1	1458	RNP2X6	RNP2X6 R.norvegicu
18	567.4	21.1	569	CQ072653	CQ072653 Sequence
19	567.4	21.1	569	CQ103466	CQ103466 Sequence

C 20	567.4	21.1	569	6	CQ142302	Sequence
C 21	567.4	21.1	569	6	CQ177794	Sequence
C 22	567.4	21.1	569	6	CQ225557	Sequence
C 23	567.4	21.1	569	6	CQ263572	Sequence
C 24	567.4	21.1	569	6	CQ300643	Sequence
C 25	567.4	21.1	569	6	CQ337952	Sequence
C 26	449.6	16.7	175167	2	AC116044	Papio ham
C 27	377	14.0	577	6	CQ073194	Sequence
C 28	377	14.0	577	6	CQ104037	Sequence
C 29	377	14.0	577	6	CQ142803	Sequence
C 30	377	14.0	577	6	CQ178299	Sequence
C 31	377	14.0	577	6	CQ226029	Sequence
C 32	377	14.0	577	6	CQ264110	Sequence
C 33	377	14.0	577	6	CQ301199	Sequence
C 34	377	14.0	577	6	CQ338447	Sequence
C 35	296.2	11.0	1558	10	RNP2X5PUR	X97328 R.norvegicu
C 36	296.2	11.0	2436	10	RNRNAP2X5	X92069 R.norvegicu
C 37	285.6	10.6	2299	10	AF333331	Mus muscu
C 38	282.6	10.5	4308	5	AF205066	Gallus ga
C 39	243.6	9.0	1978	9	AF016709	Homo sapi
C 40	240.8	8.9	1762	9	BD007172	Human P2x
C 41	240.4	8.9	2062	9	BC039015	Homo sapi
C 42	239.2	8.9	1389	9	HSP2X4PC	Y07684 H.sapiens m
C 43	239.2	8.9	1750	6	AR270268	Sequence
C 44	239.2	8.9	1762	6	AR026669	Sequence
C 45	239.2	8.9	3090	9	AF191093	Homo sapi

## ALIGNMENTS

RESULT 1	BC033488	2738 bp	mRNA	linear	PRI 25-JUN-2004
LOCUS	BC033488				
DEFINITION	Homo sapiens purinergic receptor P2X-like 1, orphan receptor, mRNA (CDNA clone IMAGE:5176046), partial cds.				
ACCESSION	BC033488				
VERSION	BC033488.1 GI:23959072				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2738)				
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Xsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., R.A., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sahegyi, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Snailus, D.E., Scherch, A., Schein, J.S., Jones, S.J., and Marra, M.A., 2001				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 2738)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (26-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				

REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: Life Technologies, Inc. cDNA Library preparation: Life Technologies cDNA Library Arrayed by: The I.M.A.G.E. Center DNA Sequencing by: Baylor College of Medicine Sequencing Center Center code: BCM-HGSC Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a> Contact: <a href="mailto:amg@bcm.tmc.edu">amg@bcm.tmc.edu</a> Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyar, Kowals, C.R., Sneed, A.J., Martin, R.G., Muzny, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 51 Row: m Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA q1: 4885534.

**FEATURES**  
**SOURCE**

```

1. 2738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:517604s"
tissue_type="brain, lung,
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
note="Vector: pCMV-SPORTS"

```

## ORIGIN

Query Match	96.0%; Score 2585.2; DB 9; Length 2738;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 2587; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	97 GTGGGCTCTCTCGCCCAAAAAGGCTACGAGGCGGGACCTGGNACCCGAGTTTTCAT 156
Db	
QY	149 GTGGGCTCTCTCGCCCAAAAAGGCTACGAGGCGGGACCTGGNACCCGAGTTTTCAT 208
Db	
QY	157 CATCACCAAACTCAAAGGGGTTTCCGCTCACTCAGATCAAGGAGCTTGGAAAACCGGCTGTG 216
Db	
QY	209 CATCACCAAACTCAAAGGGGTTTCCGCTCACTCAGATCAAGGAGCTTGGAAAACCGGCTGTG 268
Db	
QY	217 GGATGTGGCGCATCTTCGTGAAGCCACTCAGGGAGAGAACTGTCTTCTTGTGTGACAA 276
Db	
QY	269 GGATGTGGCGCACTTCGTGAAGCCACTCAGGGAGAGAACTGTCTTCTTGTGTGACAA 328
Db	
QY	277 CTTCCTTTGTGAGCGACGCCCAAGTTCAGGGCAGATGCCCAGAGACACCGTCCGTCACCT 336
Db	
QY	329 CTTCCTTTGTGAGCGACGCCCAAGTTCAGGGCAGATGCCCAGAGACACCGTCCGTCACCT 388
Db	
QY	337 GGCTAACTGCTGGGTGCAAGGACTGCCCCGAAGGGAGGGAGGCACACACAGCCACCG 396
Db	
QY	389 GGCTAACTGCTGGGTGCAAGGACTGCCCCGAAGGGAGGGAGGCACACACAGCCACCG 448
Db	
QY	397 TGTAAAAACAGGCGAGTGTGGTGTTCATCTGGACCCACAGGACCTGTGAGATCTGGAG 456
Db	
QY	449 TGTAAAAACAGGCGAGTGTGGTGTTCATCTGGACCCACAGGACCTGTGAGATCTGGAG 508
Db	
QY	457 TTGTTGCCCACTGAGAGTGGGCTTTGTGCCCTCGAGGCGCCCTGTGGCCAGGCGCCAGAA 516
Db	
QY	509 TTGTTGCCCACTGAGAGTGGGCTTTGTGCCCTCGAGGCGCCCTGTGGCCAGGCGCCAGAA 568
Db	
QY	517 CTTCACTACTGTTCACTAAAAACACAGTCACTTCACGAAGTTCAACTCTCTAAGTCCAA 576
Db	
QY	569 CTTCACTACTGTTCACTAAAAACACAGTCACTTCACGAAGTTCAACTCTCTAAGTCCAA 628
Db	
QY	577 TGCCTTGGAGACCTGGGACCCACCTATTTTAAAGACACTGCCGCTATGAAACCAATTCAG 636
Db	
QY	629 TGCCTTGGAGACCTGGGACCCACCTATTTTAAAGACACTGCCGCTATGAAACCAATTCAG 688
Db	
QY	637 CCCCCTACTGTCCGCTTCCGATTTGGGGACTGTGGCCCAAGGCTCGGAGGACCTTCGA 696
Db	

1769 GTGAGCTGGAGCCAAAGGCGAGAGGAGTGTATGGGGAGGGGATTGTTT 1828  
1777 CAGTCTCTCTGCTGTGATGCCCGCAGGAGTCTTAATCTAGGGAATGGGTGGAGTA 1836  
1829 CAGCTTCTCTGCTGTGATGCCCGCAGGAGTCTTAATCTAGGGAATGGGTGGAGTA 1888  
1837 GGCAGATTAATCACTCCCTATCCCGCAGGAGGCGGAGCATGTCTTTGGGCCACA 1896  
1889 GGCAGATTAATCACTCCCTATCCCGCAGGAGGCGGAGCATGTCTTTGGGCCACA 1948  
1897 CCGTCTTATGTTATGAGGACCGGTCTGTTTCCAGTGTAGCCCTTTTCCCATGAGGTCT 1956  
1949 CCGTCTTATGTTATGAGGACCGGTCTGTTTCCAGTGTAGCCCTTTTCCCATGAGGTCT 2008  
1957 GGCAGAGAGAGGAGGCGGAGGCTGTAAGTCTGTGATCATCTGTTTCTTCCAGACTT 2016  
2009 GGCAGAGAGAGGAGGCGGAGGCTGTAAGTCTGTGATCATCTGTTTCTTCCAGACTT 2068  
2017 CTATATCCCTCTCGGTAAACCCCGCAGGAGGCGGAGGCTGTTTCTTCCAGACTT 2076  
2069 CTATATCCCTCTCGGTAAACCCCGCAGGAGGCGGAGGCTGTTTCTTCCAGACTT 2128  
2077 TGAGAGCCCTGGGGTGGAGGCTGTGGAGGCTGTACATCTGAAATTCACCTTCAGTCCA 2136  
2129 TGAGAGCCCTGGGGTGGAGGCTGTGGAGGCTGTACATCTGAAATTCACCTTCAGTCCA 2188  
2137 AGTCATACCTAGGAAGCTGTCTGGGAGCTGTCTGAGGAGGCGGCTGCTGATCCAG 2196  
2189 AGTCATACCTAGGAAGCTGTCTGGGAGCTGTCTGAGGAGGCGGCTGCTGATCCAG 2248  
2197 GGTGATGAGTGGTGGAGGCTGTGGGAGCTGTCTGAGGAGGCGGCTGCTGATCCAG 2256  
2249 GGTGATGAGTGGTGGAGGCTGTGGGAGCTGTCTGAGGAGGCGGCTGCTGATCCAG 2308  
2257 TGCCAGGCTTTGAGCTGGAATCTCTCTTGGTCCAGGCGGCGGAGGAGTCTTAA 2316  
2309 TGCCAGGCTTTGAGCTGGAATCTCTCTTGGTCCAGGCGGCGGAGGAGTCTTAA 2368  
2317 GTGTCCACCCAGGAGGCGAAGGGCTGTCTTCCATGTGGGTACCTGGTATCAGGGCA 2376  
2369 GTGTCCACCCAGGAGGCGAAGGGCTGTCTTCCATGTGGGTACCTGGTATCAGGGCA 2428  
2377 AGCTGTGAGGCGGAGGCTGGGCTGTGAGCTGGGCTGACATCTAGATCACTGCCACC 2436  
2429 AGCTGTGAGGCGGAGGCTGGGCTGTGAGCTGGGCTGACATCTAGATCACTGCCACC 2488  
2437 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2496  
2489 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2548  
2497 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2556  
2549 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2608  
2557 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2616  
2609 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2668  
2617 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2676  
2669 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2728  
2677 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2686  
2729 TGAGGCTCAGTAAATGCTGGGCTCCCTGTGCTCTCTCAATCTCCAGAGCCATGTCCA 2738

RESULT 2  
CQ730055  
LOCUS  
DEFINITION  
ACCESSION  
VERSION

3554 bp DNA  
Sequence 15989 from Patent WO02068579.  
CQ730055  
CQ730055.1 GI:42302820

linear PAT 03-FEB-2004

KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
ORIGIN

Query Match 95.0%; Score 2558.8; DB 6; Length 3554;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2553; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

97 GTGGGCTCTCTCGCCAAAGGCTACCCAGGAGGCGGAGCTTGGAAACCCAGTTTCCAT 156  
180 GTGGGCTCTCTCGCCAAAGGCTACCCAGGAGGCGGAGCTTGGAAACCCAGTTTCCAT 239  
157 CATCAACAAACTCAAGGGGTTTCCCTCTACTAGATCAAGGAGCTTGGAAACCCGCTG 216  
240 CATCAACAAACTCAAGGGGTTTCCCTCTACTAGATCAAGGAGCTTGGAAACCCGCTG 299  
217 GGATGGGCGGCTCTGTGAAGCCACTCAGGAGAGAGACGTCTTCTTCTTGGTGACCAA 276  
300 GGATGGGCGGCTCTGTGAAGCCACTCAGGAGAGAGACGTCTTCTTCTTGGTGACCAA 359  
277 CTTCTTGTGACGCCGCCAAAGTTTCAAGGCGAGATCCAGAGCACCCGCTCTCCACT 336  
360 CTTCTTGTGACGCCGCCAAAGTTTCAAGGCGAGATCCAGAGCACCCGCTCTCCACT 419  
337 GGCTAACTGTGGGTGACGAGACTTGCCTCCGAGGAGGAGGAGGAGCACACAGCCACGG 396  
420 GGCTAACTGTGGGTGACGAGACTTGCCTCCGAGGAGGAGGAGGAGGAGGAGGAGG 479  
397 TGTAAAAAGGCGGAGTGTGTGTGTCAATGGGAGCCACAGGAGCTGTGAGATCTGGAG 456  
480 TGTAAAAAGGCGGAGTGTGTGTGTCAATGGGAGCCACAGGAGCTGTGAGATCTGGAG 539  
457 TTGTTGCCAGTGGAGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 516  
540 TTGTTGCCAGTGGAGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599  
517 CTTTCACTGTTCATCAAAACACAGTCACCTTCAAGCAAGTTCAACTTCTTAAGTCCAA 576  
600 CTTTCACTGTTCATCAAAACACAGTCACCTTCAAGCAAGTTCAACTTCTTAAGTCCAA 659  
577 TGCCTTGGAGACTTGGAGCCCTTATTTTAAAGCACTGCGGTATGAAACACAAATTCAG 636  
660 TGCCTTGGAGACTTGGAGCCCTTATTTTAAAGCACTGCGGTATGAAACACAAATTCAG 719  
637 CCCTTACTGTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 696  
720 CCCTTACTGTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 779  
697 GGACTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 756  
780 GGACTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 839  
757 CACGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 816  
840 CACGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899  
817 CTTGAGGAGGAGCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 876

[illegible]

QY	1957	GGGAGAGAGCAGAGGGCGGAGGGCTAAGTTGGTGATCATTTGGGTCTTTCAGGACCTT	2016
Db	2040	GGGAGAGAGCAGAGGGCGGAGGGCTAAGTTGGTGATCATTTGGGTCTTTCAGGACCTT	2099
QY	2017	CTATATCCCTCCCTGGTAAACCCCCAGCCAAACCCCTTGGAAATCTTCTCAGGCTTCC	2076
Db	2100	CTATATCCCTCCCTGGTAAACCCCCAGCCAAACCCCTTGGAAATCTTCTCAGGCTTCC	2159
QY	2077	TGAGAGCCCTGGGGGTGGGAGGTGTGGGAGGCTGTACATCTGAAATTCACATTCAGTCCA	2136
Db	2160	TGAGAGCCCTGGGGGTGGGAGGCTGTGGGAGGCTGTACATCTGAAATTCACATTCAGTCCA	2219
QY	2137	AGTCATACCTAGGAGCTGTCTGGGCAGCTGTCTCGAGGGAGGCCCTTGGCTCTGATCCCCAG	2196
Db	2220	AGTCATACCTAGGAGCTGTCTGGGCAGCTGTCTCGAGGGAGGCCCTTGGCTCTGATCCCCAG	2279
QY	2197	GCTGGATGAGTGGCTGGAGGAATGTTCCAAACAACACCCAGAGATCTCCCTCAGGC	2256
Db	2280	GCTGGATGAGTGGCTGGAGGAATGTTCCAAACAACACCCAGAGATCTCCCTCAGGC	2339
QY	2257	TGGCCAGGTTTTCGAGCTTGGAAATTCCTCTCTTGGTCCGAGGGCGGGCAGGAAATTCATAA	2316
Db	2340	TGGCCAGGTTTTCGAGCTTGGAAATTCCTCTCTTGGTCCGAGGGCGGGCAGGAAATTCATAA	2399
QY	2317	GTGTCCACCCAGGGAGGCAAGGGGCTGTCTTCCACTGTGGGTACTCGTGTGATCAGGGCA	2376
Db	2400	GTGTCCACCCAGGGAGGCAAGGGGCTGTCTTCCACTGTGGGTACTCGTGTGATCAGGGCA	2459
QY	2377	AGCTGTGGAGGCGCAGGGTGGGCTTGAGACTTGGGCTGACATCTAGAAATCACTGCGCAC	2436
Db	2460	AGCTGTGGAGGCGCAGGGTGGGCTTGAGACTTGGGCTGACATCTAGAAATCACTGCGCAC	2519
QY	2437	TGAGAGCTCAGTAAAAATGCTTGGGGTCCCTGCTGCTCTCAATCTCCAGAGGCATGTCCA	2496
Db	2520	TGAGAGCTCAGTAAAAATGCTTGGGGTCCCTGCTGCTCTCAATCTCCAGAGGCATGTCCA	2579
QY	2497	TGGGGAGGTGGGCTCTGAAAGGCGCAAGGTGGAGAGCAGGGCCCTTGAGGCTGGGTATC	2556
Db	2580	TGGGGAGGTGGGCTCTGAAAGGCGCAAGGTGGAGAGCAGGGCCCTTGAGGCTGGGTATC	2639
QY	2557	CAAGGAGGGGCGCTGTCACCTCATTTCTCTTGGGGCCAGAGGAGCTGATGTCAATGCT	2616
Db	2640	CAAGGAGGGGCGCTGTCACCTCATTTCTCTTGGGGCCAGAGGAGCTGATGTCAATGCT	2699
QY	2617	GGCAAAAGTCACGGAGTAAAGCCAGCAAGGCCACC	2651
Db	2700	GGCAAAAGTCACGGAGTAAAGCCAGCAAGGCCACC	2734

RESULT 3	TITLE
AB002058	JOURNAL
LOCUS	MEDLINE
DEFINITION	PUBMED
ACCESSION	REFERENCE
VERSION	AUTHORS
KEYWORDS	TITLE
SOURCE	JOURNAL
ORGANISM	
REFERENCE	
AUTHORS	

AB002058 3552 bp mRNA linear PRI 29-AUG-1997  
Homo sapiens mRNA for HUMAN P2XM, complete cds.  
AB002058  
AB002058.1 GI:2350846  
HUMAN P2XM.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (sites)  
Urano,T., Nishimori,H., Han,H., Furuhashi,T., Kimura,Y., Nakamura,Y.  
and Tokino,T.  
Cloning of P2XM, a novel human P2X receptor gene regulated by p53  
Cancer Res. 57 (15), 3281-3287 (1997)  
97354966  
9242461  
2 (bases 1 to 3552)  
Nakamura,Y.  
Direct Submission  
Submitted (22-MAR-1997) Yusuke Nakamura, The Inst. of Medical  
Science, The University of Tokyo, Lab. of Molecular Medicine, Human



Genome Center; 4-6-1, Shirokanedai Minato-ku, Tokyo, Minato-ku,  
Tokyo 108, Japan (E-mail:yusuke@ims.u-tokyo.ac.jp,  
Tel:81-3-5449-5372, Fax:81-3-5449-5433)  
Location/Qualifiers  
1. .3552  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/map="22q11"  
1. .3552  
/gene="HP2XM"  
46. .1341  
/gene="HP2XM"  
/codon\_start=1  
/product="HUMAN P2XM"  
/protein\_id="BAA22046.1"  
/db\_xref="GI:2350847"  
/translation="MGSPGATGGGLDYKTYKVMNRWRVGLRLLQFIVVYV  
GWALLKQGOERLEPQFSIIITKLGVSVTQIKELNRLWDVDFVKPPQGNVFL  
VTNPLVPAQVQGRCPHPSPVLANCWDECPGEGGTHSGVKTQCVVNGTHT  
CEIWSPEVSGVVPSPVLLAQNFILFKNTVTSKFNFSKSNALETWDPYFKHC  
RYEPFSPYCPVFRIGDLVAKAGGTFEDLALLGSGVGIHVHDCDLDTGSGGWPHYS  
FQLQKSYNPTATHWQEPGEARTLLKLYGIRFDLIVTGQAKFGLIPIAVTLGTG  
AAWLVGVTFCDLLLYVDREAHFYRTKYEEAKAPKATANSVWRELALASQARLABC  
LRRSSAPATATAAGSQTPGWPSPSSDTHLPHSGSL"

ORIGIN

Query Match	93.5%;	Score 2517.2;	DB 9;	Length 3552;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 2540;	Conservative 0;	Mismatches 13;	Indels 2;	Gaps 2;
QY	97	GTGGGCTCTCTCGCAAAAAGGCTACAGAGCGGGACCTGGAAACCCAGTTTCCAT	156	
Db	180	GTGGGCGTCTCGCCAAAAGGCTACAGAGCGGGACCTGGAAACCCAGTTTCCAT	239	
QY	157	CATCACCAAACTCAAAGGGTTTCGTCACTCAGATCAAGAGCTTGGAAACCGGCTGTG	216	
Db	240	CATCACCAAACTCAAAGGGTTTCGTCACTCAGATCAAGAGCTTGGAAACCGGCTGTG	299	
QY	217	GGATGTGCGCACTTCGTGAAGCCACCTCAGGAGAGAACCTGTCTTCTTGGTGACCAA	276	
Db	300	GGATGTGCGCACTTCGTGAAGCCACCTCAGGAGAGAACCTGTCTTCTTGGTGACCAA	359	
QY	277	CTTCTTGTGACGCGCAGCCAAAGTTCAAGGAGATGCCAGAGACCGGTCGTCCTCACT	336	
Db	360	CTTCTTGTGACGCGCAGCCAAAGTTCAAGGAGATGCCAGAGACCGGTCGTCCTCACT	419	
QY	337	GGCTAACTGCTGGTCAAGGAGTCCCGAAGGGAGGAGGACACACACGCCACGG	396	
Db	420	GGCTAACTGCTGGTCAAGGAGTCCCGAAGGGAGGAGGACACACACGCCACGG	479	
QY	397	TGTAAAAACAGCCAGTGTGTGTTTCAATGGGACCCACAGGACCTGTGAGATCTGGAG	456	
Db	480	TGTAAAAACAGCCAGTGTGTGTTTCAATGGGACCCACAGGACCTGTGAGATCTGGAG	539	
QY	457	TTGGTGCCCAAGTGGAGTGGGCTTGGCCCTCGAGGCGCCCTGTGGCCCAAGGCCAGAA	516	
Db	540	TTGGTGCCCAAGTGGAGTGGGCTTGGCCCTCGAGGCGCCCTGTGGCCCAAGGCCAGAA	599	
QY	517	CTTCACTGTTTCATCAAAAACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTCCAA	576	
Db	600	CTTCACTGTTTCATCAAAAACACAGTCACTTCAGCAAGTTCAACTTCTCTAAGTCCAA	659	
QY	577	TGCCTTGGAGACTGGGACCCCACTATTTTAAGCACTGCCGCTATGAACCAATTCAG	636	
Db	660	TGCCTTGGAGACTGGGACCCCACTATTTTAAGCACTGCCGCTATGAACCAATTCAG	719	
QY	637	CCCTACTGTCCCGTTCGCAATTTGGGACCTCGTGGCCAGGCTGGAGGACCTTCGA	696	
Db	720	CCCTACTGTCCCGTTCGCAATTTGGGACCTCGTGGCCAGGCTGGAGGACCTTCGA	779	
QY	697	GGACCTGGCGTTGCTGGGTGGCTCTGTAGGCATCAGAGTTCACTGGGATTTGTGACCTGGA	756	

Db	780	GGACCTGGCGTTGCTGGGTGGCTCTGTAGGCATCAGAGTTCACTGGGATTTGTGACCTGGA	839	
QY	757	CAACGGGAGTCTGTGCTGTGCTGCTCACTACTCCTTCAGCTGCAGGAGAGAGTACAA	816	
Db	840	CAACGGGAGTCTGTGCTGTGCTGCTCACTACTCCTTCAGCTGCAGGAGAGAGTACAA	899	
QY	817	CTTCAGGACAGCACTCACTGTGTGGAGCAACCGGGTGTGGAGGCGCCGACCTCTCAA	876	
Db	900	CTTCAGGACAGCACTCACTGTGTGGAGCAACCGGGTGTGGAGGCGCCGACCTCTCAA	959	
QY	877	GCTCTATGAATCCGCTTCGACATCTCTGTACACCGGAGAGGAGAGTTCGGGCTCAT	936	
Db	960	GCTCTATGAATCCGCTTCGACATCTCTGTACACCGGAGAGGAGAGTTCGGGCTCAT	1019	
QY	937	CCCCCGGCGCTCACACTGGGACCGGGACGCTTGGCTGGGCGTGGTCACTTTTCTG	996	
Db	1020	CCCCCGGCGCTCACACTGGGACCGGGACGCTTGGCTGGGCGTGGTCACTTTTCTG	1079	
QY	997	TGACCTGCTACTGCTGTATGTGATAGAGAACCCATTTCTACTGGAGACAAAGTATGA	1056	
Db	1080	TGACCTGCTACTGCTGTATGTGATAGAGAACCCATTTCTACTGGAGACAAAGTATGA	1139	
QY	1057	GGAGGCAAGGCGCCGAAAGCAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC	1116	
Db	1140	GGAGGCAAGGCGCCGAAAGCAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC	1199	
QY	1117	CAAGCCCGACTGGCGAGTGCCTCAGACGGAGCTCAGCACCTGCACCCACCGCCACTGC	1176	
Db	1200	CAAGCCCGACTGGCGAGTGCCTCAGACGGAGCTCAGCACCTGCACCCACCGCCACTGC	1259	
QY	1177	TGCTGGAGTCAAGACACAGACACGAGATGGCCCTGTCCAAGTCTGACACCCACTTGGC	1236	
Db	1260	TGCTGGAGTCAAGACACAGACACGAGATGGCCCTGTCCAAGTCTGACACCCACTTGGC	1319	
QY	1237	ATCCCAATTCGGGAGCTGTAGCCGTTCCCTGCTGTGTGAGAGTTGGGGGCTGGGAAGG	1296	
Db	1320	ATCCCAATTCGGGAGCTGTAGCCGTTCCCTGCTGTGTGAGAGTTGGGGGCTGGGAAGG	1379	
QY	1297	CGGGGCGCTGCTCGTGGGATCTCAAGATGAGGCCCGCAGCATGGAGATTTGGGGTAGAAT	1356	
Db	1380	CGGGGCGCTGCTCGTGGGATTTCAAGATGAGGCCCGCAGCATGGAGATTTGGGGTAGAAT	1439	
QY	1357	TCACACCTTTGMAACCCAGCAGACAGTCCCTCCCTGACTCCACCTTGTGTAGGCTGCTGC	1416	
Db	1440	TCACACCTTTGAAACCCAGCAGACAGTCCCTCCCTGACTCCACCTTGTGTAGGCTGCTGC	1499	
QY	1417	CTCAGGAGCCATAGAAAGTGGGCTGTGTTTGTGAGACGGGACAGAACCTGACCCGTTGAG	1476	
Db	1500	CTCAGGAGCCATAGAAAGTGGGCTGTGTTTGTGAGACGGGACAGAACCTGACCCGTTGAG	1559	
QY	1477	ACTGGAGAGCCAGCAGGACCTGTATTGAGGGCTCCGACTGATGTGGAGGGGCTC	1536	
Db	1560	ACTGGAGAGCCAGCAGGACCTGTATTGAGGGGCTCCGACTGATGTGGAGGGGCTC	1619	
QY	1537	CTGCTGCTGCTGGGCTGGAGGCTCTCTCCAGTGTCTGTCCCGAGTGTCTCTACATGAG	1596	
Db	1620	CTGCTGCTGCTGGGCTGGAGGCTCTCTCTCCAGTGTCTGTCTCCCGAGTGTCTCTACATGAG	1679	
QY	1597	AGGTATGCTTACAGTGTGACACAGACCTCTCTGCTGGCTGGGCTCTGGCCCTCTCTC	1656	
Db	1680	AGGTATGCTTACAGTGTGACACAGACCTCTCTGCTGGCTGGGCTCTGGCCCTCTCTC	1739	
QY	1657	CCCATCTGACCCCATCATAGGTAGAGACCCACCCCTCCCATCGGTCTCTACATGGGCT	1716	
Db	1740	CCCATCTGACCCCATCATAGGTAGAGACCCACCCCTCCCATCGGTCTCTACATGGGCT	1799	
QY	1717	GTGAGCTGGAGCAAAAAGGCAAGGAGAGAGTATGGGGAGGGGAGTGTGT	1776	
Db	1800	GTGAGCTGGAGCAAAAAGGCAAGGAGAGTATGGGGAGGGGAGTGTGT	1859	
QY	1777	CAGTTCTCTGGTGTGATGCCAGGAGTCTCTAATCTAGGAAATGGGGTGGAGTA	1836	



/db\_xref="GI:40675308"  
/db\_xref="locusID:9127"  
/db\_xref="MIM:608077"  
/translation="MGSPGATGWGLLDYKTEK WALLAKKYQERDLEPQFSITKLK  
GVSVQIKELGNRLMDVADFVKPPQV"  
110..259  
misc\_feature  
/gene="P2RXL1"  
/note="P2X\_receptor; Region: ATP P2X receptor"  
/db\_xref="CDD:pfam00864"

## ORIGIN

Query Match 88.5%; Score 2382.8; DB 9; Length 2578;

Best Local Similarity 94.4%; Pred. No. 0;

Matches 2542; Conservative 0; Mismatches 2; Indels 148; Gaps 1;

QY	2	TGCTGACTCATGTGCCCCAGCTAGCAGGAGCTGCAGCATGGGCTCCCCAGGGGCTACG	61
DB	14	TGCTGACTCATGTGCCCCAGCTAGCAGGAGCTGCAGCATGGGCTCCCCAGGGGCTACG	73
QY	62	ACAGGCTGGGGCTTCTGGATTATAGACGGAGAGTGGGCTCTCTCGCCCAAAAAGGC	121
DB	74	ACAGGCTGGGGCTTCTGGATTATAGACGGAGAGTGGGCTCTCTCGCCCAAAAAGGC	133
QY	122	TACAGGAGCGGAGCTCGAACCCCACTTTTCCATCATCAACAACTCAAAAGGGTTTCC	181
DB	134	TACAGGAGCGGAGCTCGAACCCCACTTTTCCATCATCAACAACTCAAAAGGGTTTCC	193
QY	182	CTCACTCAGATCAAGAGCTTGGAAACCGGCTGTGGGATGTGGCGACTTCGTGAAGCCA	241
DB	194	CTCACTCAGATCAAGAGCTTGGAAACCGGCTGTGGGATGTGGCGACTTCGTGAAGCCA	253
QY	242	CTTCAAGGAGAGAACGCTGTCTTCTTGTGTGACCAACTTCCCTTGTGACCCAGCCCAAGTT	301
DB	254	CTTCA-----	258
QY	302	CAGGSCAGATGCCAGAGACACCGCTCCCTCCACTGGCTAACTGCTGGGTGCAGAGGAC	361
DB	259	-----	258
QY	362	TGCCCGAAGGGAGGGAGGCACACACAGCCAGCGTGTAAAAACAGGCGAGTGTGGTG	421
DB	259	-----GGTGTAAAAACAGGCGAGTGTGGTG	285
QY	422	TTCAATGGGACCCACAGGACCTGTGAGATCTGGAGTTGGTCCCACTGGAGAGTGGGTT	481
DB	286	TTCAATGGGACCCACAGGACCTGTGAGATCTGGAGTTGGTCCCACTGGAGAGTGGGTT	345
QY	482	GTGCCCTCGAGGCCCTCTGSCCCAGCCAGACCTTCACTGTTTCATCAAAAACACA	541
DB	346	GTGCCCTCGAGGCCCTCTGSCCCAGCCAGACCTTCACTGTTTCATCAAAAACACA	405
QY	542	GTCACTTTCAGCAAGTTCAACTTCTTAAGTCCCAATGCTTGGAGACCTGGGACCCCAAC	601
DB	406	GTCACTTTCAGCAAGTTCAACTTCTTAAGTCCCAATGCTTGGAGACCTGGGACCCCAAC	465
QY	602	TATTTTAGCACTGCGGTATGAACCAATTCAGCCCTACTGTCCCGTGTTCGGCAT	661
DB	466	TATTTTAGCACTGCGGTATGAACCAATTCAGCCCTACTGTCCCGTGTTCGGCAT	525
QY	662	GGGACCTCTGTGGCCAAAGCTGGAGGACCTTTCAGGACCTGGGCTTCTGGGTGGCTCT	721
DB	526	GGGACCTCTGTGGCCAAAGCTGGAGGACCTTTCAGGACCTGGGCTTCTGGGTGGCTCT	585
QY	722	GTAGGCATCAGAGTTCACTGGGATTTGACCTTGGACACCGGGGACTCTGGCTGTGGCT	781
DB	586	GTAGGCATCAGAGTTCACTGGGATTTGACCTTGGACACCGGGGACTCTGGCTGTGGCT	645
QY	782	CACTACTCTTCCAGCTCAGAGAGAGCTACACTTCAGACAGCCACTCACTGGTGG	841
DB	646	CACTACTCTTCCAGCTCAGAGAGAGCTACACTTCAGACAGCCACTCACTGGTGG	705
QY	842	GAGCAACCGGGTGTGGAGCCCGCACTCTGCTCAAGCTCTATGGAATCCGCTTCGACATC	901

DB	706	GAGCAACCGGTTGTGGAGGCGCCGACCCCTGCTCAAGCTCTATGGAATCCGCTTCGACATC	765
QY	902	CTCGTCAACCGGGCAGCAGGAGAGTTTCGGGCTCATCCCCAGGCGCTCACTCTGGGACCC	961
DB	766	CTCGTCAACCGGGCAGCAGGAGAGTTTCGGGCTCATCCCCAGGCGCTCACTCTGGGACCC	825
QY	962	GGGSCAGCTTGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1021
DB	826	GGGSCAGCTTGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	885
QY	1022	AGAGAGAGCCATTTCTACTGGAGGACAAAGTATAGAGGCGCAAGGCCCGGCAAGCAAC	1081
DB	886	AGAGAGAGCCATTTCTACTGGAGGACAAAGTATAGAGGCGCAAGGCCCGGCAAGCAAC	945
QY	1082	GCCAACTCTGTGTGGAGGAGCTGGCCCTTGATCCCAAGCCGAGCTGCGGAGTGCCTC	1141
DB	946	GCCAACTCTGTGTGGAGGAGCTGGCCCTTGATCCCAAGCCGAGCTGCGGAGTGCCTC	1005
QY	1142	AGACGGAGCTCAGACACTGCACCCAGCGCACTGTGCTGGGAGTCAGACACAGACCA	1201
DB	1006	AGACGGAGCTCAGACACTGCACCCAGCGCACTGTGCTGGGAGTCAGACACAGACCA	1065
QY	1202	GGATGGCCCTGTCCAAAGTCTGACACCCACTTGCACAAACCCTTCCGGGAGCTGTAGCG	1261
DB	1066	GGATGGCCCTGTCCAAAGTCTGACACCCACTTGCACAAACCCTTCCGGGAGCTGTAGCG	1125
QY	1262	TTCCCTGCTGCTGAGAGTTGGGGCTGGAGGGGGGGGGCTGCTGCTGGGATCTCAAG	1321
DB	1126	TTCCCTGCTGCTGAGAGTTGGGGCTGGAGGGGGGGGGCTGCTGCTGGGATCTCAAG	1185
QY	1322	GATGAGGCGCCAGCATGGAGGATTTGGGGTGAATTTCAACCTTGAACCCAGCAGACAG	1381
DB	1186	GATGAGGCGCCAGCATGGAGGATTTGGGGTGAATTTCAACCTTGAACCCAGCAGACAG	1245
QY	1382	TCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1441
DB	1246	TCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1305
QY	1442	TGTTTTGAGACGGCGCAGAACCTGACCCGCTGGAGACTGGGAGAGCCAGCAGGACCTG	1501
DB	1306	TGTTTTGAGACGGCGCAGAACCTGACCCGCTGGAGACTGGGAGAGCCAGCAGGACCTG	1365
QY	1502	TATTCAGGGCTCCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1561
DB	1366	TATTCAGGGCTCCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1425
QY	1562	CTCTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1621
DB	1426	CTCTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1485
QY	1622	AGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1681
DB	1486	AGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1545
QY	1682	GAGACCCCACTCCCATCGGCTCTACATGGGGCTGTCAGCTGGAGCCCAAAAGGCAAG	1741
DB	1546	GAGACCCCACTCCCATCGGCTCTACATGGGGCTGTCAGCTGGAGCCCAAAAGGCAAG	1605
QY	1742	GCAGAAAGGAGGAGTGTATGGGGAGGGGATTTGTTTTCAGCTTCTCTGCTGCTGCTGCTG	1801
DB	1606	GCAGAAAGGAGGAGTGTATGGGGAGGGGATTTGTTTTCAGCTTCTCTGCTGCTGCTGCTG	1665
QY	1802	CAGAGAGTCTTAATCTAGGGAATGGGTGGAGTAGGCAGATTAATCCACTCCCTATCCC	1861
DB	1666	CAGAGAGTCTTAATCTAGGGAATGGGTGGAGTAGGCAGATTAATCCACTCCCTATCCC	1725
QY	1862	CCAGGCAAGGGCGGAGCATGTCTTTGGGCGCCACACTGCTGTTAGTTTATGAGGACCGGCT	1921
DB	1726	CCAGGCAAGGGCGGAGCATGTCTTTGGGCGCCACACTGCTGTTAGTTTATGAGGACCGGCT	1785
QY	1922	GCCTTTCAGTGTAGCCCTTTTGCCTATGGAGTCTGGGAGAGAGAGAGGCGGCGCAG	1981
DB	1786	GCCTTTCAGTGTAGCCCTTTTGCCTATGGAGTCTGGGAGAGAGAGAGGCGGCGCAG	1845

QY 1982 GCTAGTTGGTGATCATTTGGGTTCTTCAGGACCTCTATATCCCTCCTCGGTACCCCCC 2041  
Db 1846 GCTAAGTTGGTGATCATTTGGGTTCTTCAGGACCTCTATATCCCTCCTCGGTACCCCCC 1905  
QY 2042 AGCCCAACCCCTTGGAATCTTTTCTCCAGGCTTCTCAGAGCCCTCGGGGTGGAGGCTG 2101  
Db 1906 AGCCCAACCCCTTGGAATCTTTTCTCCAGGCTTCTCAGAGCCCTCGGGGTGGAGGCTG 1965  
QY 2102 TGGGAGGCTGTACATCTCAAAATTCATTCAGTCAAGTCATACCTAGGAGCTCTCGG 2161  
Db 1966 TGGGAGGCTGTACATCTCAAAATTCATTCAGTCAAGTCATACCTAGGAGCTCTCGG 2025  
QY 2162 CAGCTGCTCGAGGAGGCGCCCTGGCTCTGATCCAGGCTGGAATGAGTGGTGAAGGAAT 2221  
Db 2026 CAGCTGCTCGAGGAGGCGCCCTGGCTCTGATCCAGGCTGGAATGAGTGGTGAAGGAAT 2085  
QY 2222 GGTTCACAAACACACACAGGAGTCTCCTCAGGCTGCGCAGGTTTTCAGCTGGAATTC 2281  
Db 2086 GGTTCACAAACACACACAGGAGTCTCCTCAGGCTGCGCAGGTTTTCAGCTGGAATTC 2145  
QY 2282 TCCTCTGCTGCTCAGGAGGCGGCGAGGAAATCTAAGTGTCCACCCAGGAGGCAAGGG 2341  
Db 2146 TCCTCTGCTGCTCAGGAGGCGGCGAGGAAATCTAAGTGTCCACCCAGGAGGCAAGGG 2205  
QY 2342 CTGCTTTCCACTGTGGTACCTGTGTATCAGGCGAAGCTGTGAGGCGCCAGGGTGGGGC 2401  
Db 2206 CTGCTTTCCACTGTGGTACCTGTGTATCAGGCGAAGCTGTGAGGCGCCAGGGTGGGGC 2265  
QY 2402 TGAGCTGGGTGACATCTGAATCAGCTGCACCTGAGCCTCAGTAAATGCCTGGG 2461  
Db 2266 TGAGCTGGGTGACATCTGAATCAGCTGCACCTGAGCCTCAGTAAATGCCTGGG 2325  
QY 2462 TCCCTGCTGCTCTCAATCTCCAGAGCATGTCCATGGGAGGTGGGCTCTGAAGGCGGA 2521  
Db 2326 TCCCTGCTGCTCTCAATCTCCAGAGCATGTCCATGGGAGGTGGGCTCTGAAGGCGGA 2385  
QY 2522 AGTGGGAGACAGGCGCCCTGAGGCTGGTATCCAGAGGCGGACGTGCACTGTATT 2581  
Db 2386 AGTGGGAGACAGGCGCCCTGAGGCTGGTATCCAGAGGCGGACGTGCACTGTATT 2445  
QY 2582 CTCCTTGGGCGCCAGAGAGCTGATGTCTGCTGCAAGTCAAGTCAAGTCAAGTCAAG 2641  
Db 2446 CTCCTTGGGCGCCAGAGAGCTGATGTCTGCTGCAAGTCAAGTCAAGTCAAGTCAAG 2505  
QY 2642 CAAGCCACCAAAAAAIAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2693  
Db 2506 CAAGCCACCAAAAAAIAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2557  
  
RESULT 5  
AC002472 147086 bp DNA linear PRI 09-APR-2003  
LOCUS Homo sapiens Chromosome 22q11.2 PAC Clone p\_n5 In BCLR2-GGT Region,  
DEFINITION complete sequence.  
AC002472  
ACCESSION AC002472.8 GI:24137490  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Budarf, M.L. and Emanuel, B.S.  
JOURNAL Unpublished  
REFERENCE  
AUTHORS Zhang, G., Zhan, M., Lao, V. and Roe, B.A.  
JOURNAL Unpublished  
TITLE Homo sapiens Chromosome 22q11.2 PAC Clone p\_n5 In BCLR2-GGT Region  
REFERENCE  
AUTHORS Zhang, G., Lao, V., Zhan, M. and Roe, B.A.  
JOURNAL Unpublished  
TITLE Direct Submission  
JOURNAL Submitted (20-AUG-1997) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
4 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (10-SEP-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
5 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (23-OCT-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
6 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (13-NOV-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
7 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (18-NOV-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
8 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (21-NOV-1998) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
9 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (14-APR-1999) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
10 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (13-SEP-1999) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
11 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (13-MAY-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
12 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
13 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (20-MAY-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
14 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (31-MAY-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
15 (bases 1 to 147086)  
Zhang, G., Zhan, M., Budarf, M.L., Emanuel, B.S. and Roe, B.A.  
Direct Submission  
Submitted (01-SEP-2002) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

REFERENCE	AUTHORS	TITLE	JOURNAL
OK 73019, USA	Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.	Submitted (19-OCT-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
REFERENCE	AUTHORS	TITLE	JOURNAL
OK 73019, USA	Zhang,G., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.	Submitted (09-APR-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
COMMENT	On Oct 19, 2002 this sequence version replaced gi.22597497. Because these overlapping clones came from different libraries.		
FEATURES	Location/Qualifiers 1..147086 <del>/organism="Homo sapiens"</del> /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="22" /map="22q11.2" /clone="P_n5"		
ORIGIN	Query Match 59.1%; Score 1592.6; DB 9; Length 147086; Best Local Similarity 99.7%; Pred. No. 0; Matches 1595; Conservative 0; Mismatches 4; Indels 0; Gaps 0;		
QY	1053	ATGAGGCGCAAGGCCCGAAGAACCCGCACACTCTGTGTGGAGGAGTGCCCTTG	1112
Db	70561	ATTGCGAGCACAAGGCCCGAAGAACCCGCACACTCTGTGTGGAGGAGTGCCCTTG	70620
QY	1113	CATCCCAAGCCGACCTGGCGAGTGCCTCAGACGAGCTCAGCACCTGCACCACGGCCA	1172
Db	70621	CATCCCAAGCCGACCTGGCGAGTGCCTCAGACGAGCTCAGCACCTGCACCACGGCCA	70680
QY	1173	CTGCTGTGGAGTCAGACACACACAGGATGCCCTGTCCAAGTTCTGACACCCACT	1232
Db	70681	CTGCTGTGGAGTCAGACACACACAGGATGCCCTGTCCAAGTTCTGACACCCACT	70740
QY	1233	TGCCAAACCATTCGGGACCTGTAGCCGTTCCCTGCTGTGTGAGTGGGGCTGGGA	1292
Db	70741	TGCCAAACCATTCGGGACCTGTAGCCGTTCCCTGCTGTGTGAGTGGGGCTGGGA	70800
QY	1293	AGGCGGGGCCCCCTGCTGGGGATCTCAAGGATAGGCGCCAGCATGGAGATTGGGGTA	1352
Db	70801	AGGCGGGGCCCCCTGCTGGGGATCTCAAGGATAGGCGCCAGCATGGAGATTGGGGTA	70860
QY	1353	GAAATCCACCTTGAACCCACACAGACAGTCCCTCCCTGATCCCACTTGTGGGTG	1412
Db	70861	GAAATCCACCTTGAACCCACACAGACAGTCCCTCCCTGATCCCACTTGTGGGTG	70920
QY	1413	CTGCTCAGGAGCCATAGAAGTCGGCTGTGTTTTGAGCGCGACAGAACCTGACCCGT	1472
Db	70921	CTGCTCAGGAGCCATAGAAGTCGGCTGTGTTTTGAGCGCGACAGAACCTGACCCGT	70980
QY	1473	GGAGACTGGAGAGCCACAGGACCTGTATGTGAGGCTCCGACTGCACTGTGGCAGGG	1532
Db	70981	GGAGACTGGAGAGCCACAGGACCTGTATGTGAGGCTCCGACTGCACTGTGGCAGGG	71040
QY	1533	GCTCTGTGCTGTGTGGGCTGGAGGTCCTCTCCAGTGTCTGTCCCGAGTGTTCCTA	1592
Db	71041	GCTCTGTGCTGTGTGGGCTGGAGGTCCTCTCCAGTGTCTGTCCCGAGTGTTCCTA	71100
QY	1593	GCAGAGGATATGCTTACAGCTGTGACGACAGACCCCTCTGTGCTGGGTCCTGGCCCTC	1652
Db	71101	GCAGAGGATATGCTTACAGCTGTGACGACAGACCCCTCTGTGCTGGGTCCTGGCCCTC	71160
QY	1653	CTCCCCAFCTGACCCCACATCATAGGTAGAGACCCCACTCCCACTCGCTCTCATGG	1712
Db	71161	CTCCCCAFCTGACCCCACATCATAGGTAGAGACCCCACTCCCACTCGCTCTCATGG	71220











Db 840 CACCGGGACTCTGGCTGCTGCTCACTACTCTTCCAGTGCAGGAGAGAGTACAA 899  
QY 817 CTTACAGACAGCCACTCACTGTGGAGCAACCGGGTGTGGAGCCCGCACCCCTGCTCA 876  
Db 900 CTTACAGACAGCCACTCACTGTGGAGCAACCGGGTGTGGAGCCCGCACCCCTGCTCA 959  
QY 877 GCTCTATGGAATCCGCTTGCACATCTCTGTCACCCGGCAGGAGGAAAGTTCGGGCTCAT 936  
Db 960 GCTCTATGGAATCCGCTTGCACATCTCTGTCACCCGGCAGGAGGAAAGTTCGGGCTCAT 1019  
QY 937 CCCACGGCCGTCACACTGGGACCGGGACGCTTGGCTGGGGGCTGACCTTTTCTG 996  
Db 1020 CCCACGGCCGTCACACTGGGACCGGGACGCTTGGCTGGGGGCTGACCTTTTCTG 1079  
QY 997 TGACCTGTACTGTCTGTATGTGATAGAGAAGCCATTTCTACTGGAGGACAAAGTATGA 1056  
Db 1080 TGACCTGTACTGTCTGTATGTGATAGAGAAGCCATTTCTACTGGAGGACAAAGTATGA 1139  
QY 1057 GGAGGCCAAGGCCCGGAAAGCAACCGCCTCTGTGTGGAGGAGCTGGGCCCTTGCATC 1116  
Db 1140 GGAGGCCAAGGCCCGGAAAGCAACCGCCTCTGTGTGGAGGAGCTGGGCCCTTGCATC 1199  
QY 1117 CCAAGCCGACTGGCCGAGTGCTCAGACGGAGCTCAGACCTGCACCCAGCGCCACTGC 1176  
Db 1200 CCAAGCCGACTGGCCGAGTGCTCAGACGGAGCTCAGACCTGCACCCAGCGCCACTGC 1259  
QY 1177 TGTGGAGTCAACACACAGACACAGGATGGCCCTGTCCAACTTCTGACACCCACTTGCC 1236  
Db 1260 TGTGGAGTCAACACACAGACACAGGATGGCCCTGTCCAACTTCTGACACCCACTTGCC 1319  
QY 1237 AACCATTCCGGAGCCTGTAGCCGTTCCCTGCTGTTGAGAGTTGGGGCTGGGAAGG 1296  
Db 1320 AACCATTCCGGAGCCTGTAGCCGTTCCCTGCTGTTGAGAGTTGGGGCTGGGAAGG 1379  
QY 1297 CGGGCCCTGCTGGGGATCTCAAGGATGAGCCGACGATGAGAGATGGGGGTAGAA 1356  
Db 1380 CGGGCCCTGCTGGGGATTTCAAGATGAGCCGACGATGAGAGATGGGGGTAGAA 1439  
QY 1357 TCACCCCTGACCCAGACAGACAGTCCCTCCCTGACTCCACCTTGGTAGGGTGTGC 1416  
Db 1440 TCACCCCTGACCCAGACAGACAGTCCCTCCCTGACTCCACCTTGGTAGGGTGTGC 1499  
QY 1417 CTCAGGAGCCATAGAAATCGGCTGTGTTTGGAGCGGCGAGAACCTGACCCCTGGAG 1476  
Db 1500 CTCAGGAGCCATAGAAATCGGCTGTGTTTGGAGCGGCGAGAACCTGACCCCTGGAG 1559  
QY 1477 ACTGGAGAGCCAGCAGCAGCTGTATGAGGCTCCGACTGCACTGTCAGAGGGCTC 1536  
Db 1560 ACTGGAGAGCCAGCAGCAGCTGTATGAGGCTCCGACTGCACTGTCAGAGGGCTC 1619  
QY 1537 CTGCTGCGTCTGGGCTGGAGTCTCTCCAGTCTGTCCTCCAGTCTGTCCTTAGCAG 1596  
Db 1620 CTGCTGCGTCTGGGCTGGAGTCTCTCTCCAGTCTGTCCTCCAGTCTGTCCTTAGCAG 1679  
QY 1597 AGGTATGCTTACAGCTG 1614  
Db 1680 AGGTATGCTTACAGCTG 1697

RESULT 9  
E17366  
LOCUS Human mRNA for P2XM protein complete cds. 1697 bp DNA linear PAT 28-JUL-1999  
DEFINITION E17366  
ACCESSION E17366  
VERSION E17366.1 GI:5712049  
KEYWORDS JP 199826281-A/2.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1697)  
AUTHORS Tokino T. and Nakamura, Y.  
TITLE HUMAN GENE

JOURNAL  
COMMENT  
Patent: JP 199826281-A 2 06-OCT-1998;  
OTSUKA PHARMACEUT CO LTD  
OS Homo sapiens (human)  
PN JP 199826281-A/2  
PD 06-OCT-1998  
PF 26-NAR-1997 JP 1997093044  
PI TOKINO TAKASHI, NAKAMURA YUSUKE  
PC C12N15/09, A61K48/00, C07H21/04, C07K16/32, C12P21/02,  
PC C12Q1/68,  
PC  
G01N33/53, (C12N15/09, C12R1:91), (C12P21/02, C12R1:91), (C12P21/02, PC  
C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
FH Key Location/Qualifiers  
FT source 1..1697  
FT CDS /organism="Homo sapiens",  
FT /mol\_type="genomic DNA",  
FT /product="P2XM protein".  
FEATURES  
source 1..1697  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 56.3%; Score 1515.4; DB 6; Length 1697;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1512; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 97 GTGGCTCTCTCGCCAAAAGGCTACAGGAGCGGACCTGGAACCCAGTTTCCAT 156  
Db 180 GTGGCGCTCTCGCCAAAAGGCTACAGGAGCGGACCTGGAACCCAGTTTCCAT 239  
QY 157 CATCAAACTCAAGGGGTTTCCCTCACTCAGTCAAGGAGCTTGGAAACCGGCTG 216  
Db 240 CATCAAACTCAAGGGGTTTCCCTCACTCAGTCAAGGAGCTTGGAAACCGGCTG 299  
QY 217 GGATGGCGGCTTCGTGAAGCCACCTCAGGAGAGAACGTGTTCTTCTGTGACCAA 276  
Db 300 GGATGGCGGCTTCGTGAAGCCACCTCAGGAGAGAACGTGTTCTTCTGTGACCAA 359  
QY 277 CTTCTTGTGAGCCGACCCCAAGTTTCAAGGAGATGCCAGAGACCCCTGCTCCACT 336  
Db 360 CTTCTTGTGAGCCGACCCCAAGTTTCAAGGAGATGCCAGAGACCCCTGCTCCACT 419  
QY 337 GGCTAACTGTGGGTGCAGGAGACTGCCCGAAGGGAGGAGGACACACAGCCACGG 396  
Db 420 GGCTAACTGTGGGTGCAGGAGACTGCCCGAAGGGAGGAGGACACACAGCCACGG 479  
QY 397 TGTAAAAAAGCCAGCTGTGTGTTCAATGGGACCCACAGACCTGTGAGATCTGGAG 456  
Db 480 TGTAAAAAAGCCAGCTGTGTGTTCAATGGGACCCACAGACCTGTGAGATCTGGAG 539  
QY 457 TTGTGCCAGTGGAGTGGCGTGTGCCCTGAGGCCCTGCTGGCCAGGCCAGAA 516  
Db 540 TTGTGCCAGTGGAGTGGCGTGTGCCCTGAGGCCCTGCTGGCCAGGCCAGAA 599  
QY 517 CTTACACTCTTCAATAAAACACAGTCACTTTCAGCAAGTTCAACTTCTTAAGTCCAA 576  
Db 600 CTTACACTCTTCAATAAAACACAGTCACTTTCAGCAAGTTCAACTTCTTAAGTCCAA 659  
QY 577 TGGCTTGGAGACCTGGGACCCACCTATTTAAGCACTGCCGTATGAACCAATTCAG 636  
Db 660 TGGCTTGGAGACCTGGGACCCACCTATTTAAGCACTGCCGTATGAACCAATTCAG 719  
QY 637 CCCTACTGTCCCGTGTTCGCAATTGGGACCTCTGTCGCAAGCTGGAGGACCTTCGA 696  
Db 720 CCCTACTGTCCCGTGTTCGCAATTGGGACCTCTGTCGCAAGCTGGAGGACCTTCGA 779  
QY 697 GGACCTGGCGCTTCTGCGGTCTCTGTAGGATCAGAGTTCACTGGGATGTGACCTGGA 756

Db 780 GGACCTGGCTTGTGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTTGACCTGGA 839  
QY 757 CACCGGGACTCTGGCTGGCTCTACTCTCTTCCAGTGCAGGAGAGCTACAA 816  
Db 840 CACCGGGACTCTGGCTGGCTCTACTCTCTTCCAGTGCAGGAGAGCTACAA 899  
QY 817 CTTGAGGACCGCACTCACTGGTGGGAGCAACCGGGTGTGGAGCCCGCACTGCTCAA 876  
Db 900 CTTGAGGACCGCACTCACTGGTGGGAGCAACCGGGTGTGGAGCCCGCACTGCTCAA 959  
QY 877 GCTCTATGGAATCCGCTTCGACATCCCTGCTCACCGGCGAGCAGGGAAGTTGCGGCTCAT 936  
Db 960 GCTCTATGGAATCCGCTTCGACATCCCTGCTCACCGGCGAGCAGGGAAGTTGCGGCTCAT 1019  
QY 937 CCCACGCGCTCACACTGGGACCGGGGAGCTTGGCTGGGCTGCTCACTTTTCTG 996  
Db 1020 CCCACGCGCTCACACTGGGACCGGGGAGCTTGGCTGGGCTGCTCACTTTTCTG 1079  
QY 997 TGACCTGCTACTGCTGTATGTGGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1056  
Db 1080 TGACCTGCTACTGCTGTATGTGGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1139  
QY 1057 GGAGCCCAAGCCCGGAAAGCAACCGCACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1116  
Db 1140 GGAGCCCAAGCCCGGAAAGCAACCGCACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1199  
QY 1117 CCAGCCCGACTGGCCGAGTGCCTCAGACGGAGCTCAGCACTGCACCCACGGCCACTGC 1176  
Db 1200 CCAGCCCGACTGGCCGAGTGCCTCAGACGGAGCTCAGCACTGCACCCACGGCCACTGC 1259  
QY 1177 TGCTGGGAGTCAGACACAGACACACAGGATGCCCTGTGCCAAGTTCTGACACCACTTGCC 1236  
Db 1260 TGCTGGGAGTCAGACACAGACACACAGGATGCCCTGTGCCAAGTTCTGACACCACTTGCC 1319  
QY 1237 AACCATTTCCGGAGCTGTAGCGTTTCCCTGCTGGTTGAGAGTTGGGCTGGAGGG 1296  
Db 1320 AACCATTTCCGGAGCTGTAGCGTTTCCCTGCTGGTTGAGAGTTGGGCTGGAGGG 1379  
QY 1297 CCGGCCCCCTGCTGGGATCTCAAGGATGAGGCCGCCAGCATGGAGGATTTGGGGTAGAAT 1356  
Db 1380 CCGGCCCCCTGCTGGGATTTCAAGGATGAGGCCGCCAGCATGGAGGATTTGGGGTAGAAT 1439  
QY 1357 TCCACCTTTGAACCCAGCAGACAGTCCCTCCCTGACTCCACCTTGTGTAGGTGTGTC 1416  
Db 1440 TCCACCTTTGAACCCAGCAGACAGTCCCTCCCTGACTCCACCTTGTGTAGGTGTGTC 1499  
QY 1417 CTCAGGAGCCATAGAAGTCGGCTGTGTTTGTAGACGGCGACAGAACTGACCCGTTGGAG 1476  
Db 1500 CTCAGGAGCCATAGAAGTCGGCTGTGTTTGTAGACGGCGACAGAACTGACCCGTTGGAG 1559  
QY 1477 ACTGGAGAGCCAGCAGGACCTGATATGAGGGCTCCGACTGCATGTGGCAGGGCTC 1536  
Db 1560 ACTGGAGAGCCAGCAGGACCTGATATGAGGGCTCCGACTGCATGTGGCAGGGCTC 1619  
QY 1537 CTGCTGCTTGGGCTGGAGTCTCTCTCCAGTGTCTGTCCCGAGTGTCTCTAGCAG 1596  
Db 1620 CTGCTGCTTGGGCTGGAGTCTCTCTCCAGTGTCTGTCTCCCGAGTGTCTCTAGCAG 1679  
QY 1597 AGGTATGCTTACCAGCTG 1614  
Db 1680 AGGTATGCTTACCAGCTG 1697

RESULT 10  
CR456535 1517 bp mRNA linear PRI 25-MAY-2004  
LOCUS  
DEFINITION  
CR456535 Homo sapiens P2RXL1 full length open reading frame (ORF) cDNA clone  
(cDNA clone C22ORF.pGEM.P2RXL1).  
ACCESSION  
CR456535  
VERSION  
CR456535.1 GI:47678600  
KEYWORDS  
CDNA; chromosome 22; ORF.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1517)  
Collins, J.E., Wright, C.L., Edwards, C.A., Davis, M.P., Grinham, J.A.,  
Cole, C.G., Goward, M.E., Aguado, B., Mallia, M., Mokrab, Y.,  
Huckle, E.J., Beare, D.M. and Dunham, I.  
Direct Submission  
Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: c22@sanger.ac.uk Manuscript  
Sanger Institute name: PGEM.P2RXL1  
Homo sapiens cDNA sequence. This sequence was generated as part of  
the Wellcome Trust Sanger Institute program to isolate cDNA clones  
representing the full length open reading frame of well annotated  
protein coding genes on human chromosome 22. For more information  
see <http://www.sanger.ac.uk/HGP/Chr22/>.  
Location/Qualifiers  
1..1517  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clone="pGEM.P2RXL1"  
/lab\_host="JM109"  
1..1517  
/gene="P2RXL1"  
7..1332  
/gene="P2RXL1"  
/codon\_start=1  
/protein\_id="CAAG30421.1"  
/db\_xref="GI:47678601"  
/translation="MCPOLAGAGSMGSPGATGWLDDYKTEKYVMENRVRVGLORL  
IQFGIVYVGNWALLAKGYQERDLEPOFSIITKLKGVSVTQIKELGNRLNDVADFVK  
PPQENVFVLTNLTVAQVQRCPEHPSPLANCWDEDCPEBGGTHSHGVKIQG  
CVFNGHRTCEIISWCVESGVPSFLAQAFNFTLFIKNTVTFKFNFSKNALE  
TWDPYFKHCRYEPOFSYCFVTRIGDLVAKAGGTFEDALLGSGVIRVHWDCLDIT  
GDSGCMHYSPLOESKSYNFTATHWSPGVSEARTLLKYGIRFDILVTCQAKFGFL  
IPTATLCTGAAWGVTFPCDLLLYVDREAHFVWRTKYBEAKAPKATANSVRELA  
LASQRLAECLRRSSAPAPTATAGSQITQFWPCPSSDTHLPHSGSL"  
Query Match 49.6%; Score 1336.8; DB 9; Length 1517;  
Best Local Similarity 99.8%; Pred. No. 9.2e-312;  
Matches 1345; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 97 GTGGGCTCTCTCGCAAAAAGGCTACAGGAGGGGACCTGGAAACCCAGTTTCCAT 156  
Db 171 GTGGGCGCTCTCTCGCAAAAAGGCTACAGGAGGGGACCTGGAAACCCAGTTTCCAT 230  
QY 157 CATCACCAAACTCAAAAGGGTTTCCTGCTCACTCAGATCAAGGAGCTTGGAAACCCGCTGTG 216  
Db 231 CATCACCAAACTCAAAAGGGTTTCCTGCTCACTCAGATCAAGGAGCTTGGAAACCCGCTGTG 290  
QY 217 GGATGTGGCCGACTTCTGTGAAGCCACCTCAGGAGAGAAACGTGTCTTCTTGGTGACCAA 276  
Db 291 GGATGTGGCCGACTTCTGTGAAGCCACCTCAGGAGAGAAACGTGTCTTCTTGGTGACCAA 350  
QY 277 CTTCTCTGTGCGCCAGCCAGCCAGTTTCAGGCGAGATGCCAGAGACCCGCTCGTCCCACT 336  
Db 351 CTTCTCTGTGCGCCAGCCAGCCAGTTTCAGGCGAGATGCCAGAGACCCGCTCGTCCCACT 410  
QY 337 GGCTAACTCTGGGTGAGCGAGGACTGCCCGAAGGGAGGAGGACACACAGCCAGCG 396  
Db 411 GGCTAACTCTGGGTGAGCGAGGACTGCCCGAAGGGAGGAGGACACACAGCCAGCG 470  
QY 397 TGTAATAACAGCCAGTGTGTGTTTCAATGGGACCCACAGGACCTGTGATCTGGAG 456  
Db 471 TGTAATAACAGCCAGTGTGTGTTTCAATGGGACCCACAGGACCTGTGATCTGGAG 530  
QY 457 TTGGTGCCCGCAGTGGAGTGGCGTTGTGCGCTTCGAGGCGCCCTGTGTGGCCCGAGCAA 516  
Db 531 TTGGTGCCCGCAGTGGAGTGGCGTTGTGCGCTTCGAGGCGCCCTGTGTGGCCCGAGCAA 590  
QY 517 CTTTACACTGTTTCATCAAAAACACAGTCACCTTTCAGCAAGTTCACTTCTCTAAGTCCAA 576

Db	591	CTTCACACTGTTTCATCAAAACACAGTCACCTTCAGCAAGTTCACACTTCTCTAAGTCCAA	650	
QY	577	TGCCTTGGAGACTGGAGACCCCACTATTTTAAGCACTGCCGTATGAAACCAATTCAG	636	
Db	651	TGCCTTGGAGACTGGAGACCCCACTATTTTAAGCACTGCCGTATGAAACCAATTCAG	710	
QY	637	CCCCTACTGTCGCCGTGTTCCGATTTGGGACCTCGTGCCCAAGGCTGGAGGGACCTTCGA	696	
Db	711	CCCCTACTGTCGCCGTGTTCCGATTTGGGACCTCGTGCCCAAGGCTGGAGGGACCTTCGA	770	
QY	697	GGACCTGGCGTTGCTGGGTGCTCTGTAGGCATCAGAGTTCACTGGGATTTGTGACCTGGA	756	
Db	771	GGACCTGGCGTTGCTGGGTGCTCTGTAGGCATCAGAGTTCACTGGGATTTGTGACCTGGA	830	
QY	757	CACCGGGACTCTGCGTCTGGCCCTCACTACTCCTTCCAGCTGCAGAGAGAGCTACAA	816	
Db	831	CACCGGGACTCTGCGTCTGGCCCTCACTACTCCTTCCAGCTGCAGAGAGAGCTACAA	890	
QY	817	CTTCAGGACAGCCACTCACTGGTGGGAGCAACCGGGTGTGGAGGCCCGCACCTTGCTCAA	876	
Db	891	CTTCAGGACAGCCACTCACTGGTGGGAGCAACCGGGTGTGGAGGCCCGCACCTTGCTCAA	950	
QY	877	GCTCTATGGAATCCGCTTCGACATCTCTGTCACCGGCGAGGAGGAAAGTTCCGGCTCAT	936	
Db	951	GCTCTATGGAATCCGCTTCGACATCTCTGTCACCGGCGAGGAGGAAAGTTCCGGCTCAT	1010	
QY	937	CCCCACGGCGCTCACACTGGGCACCGGGGAGCTTGGCTGGCGGTGTCACTTTTCTG	996	
Db	1011	CCCCACGGCGCTCACACTGGGCACCGGGGAGCTTGGCTGGCGGTGTCACTTTTCTG	1070	
QY	997	TGACCTGCTACTGCTGTATGTGGATAGAGAAGCCCATTTCTACTGGAGGACAAAGTATGA	1056	
Db	1071	TGACCTGCTACTGCTGTATGTGGATAGAGAAGCCCATTTCTACTGGAGGACAAAGTATGA	1130	
QY	1057	GGAGGCCAAGGCCCGAAAGCAACCGCAACTCTGTGTGAGGAGAGCTGGCCCTTGCATC	1116	
Db	1131	GGAGGCCAAGGCCCGAAAGCAACCGCAACTCTGTGTGAGGAGAGCTGGCCCTTGCATC	1190	
QY	1117	CCAAGCCGAGCTGGCCGAGTGCCTCAGACGGAGCTCAGACCTGCACCCACGCGCACTGC	1176	
Db	1191	CCAAGCCGAGCTGGCCGAGTGCCTCAGACGGAGCTCAGACCTGCACCCACGCGCACTGC	1250	
QY	1177	TGCTGGGAGTCAGACACAGACACACAGGATGGCCCTGTCCAAGTTCTGACACCCACTTGC	1236	
Db	1251	TGCTGGGAGTCAGACACAGACACACAGGATGGCCCTGTCCAAGTTCTGACACCCACTTGC	1310	
QY	1237	AACCCATTCGGGAGCCTGTAGCGGTTCCCTGCTGTGTGAGAGTTGGGGCTGGGAAGG	1296	
Db	1311	AACCCATTCGGGAGCCTGTAGCGGTTCCCTGCTGTGTGAGAGTTGGGGCTGGGAAGG	1370	
QY	1297	CGGGGCCCTGCTGGGGATCTCAAGGATGAGGCCCGCCAGCATGGAGATTGGGGGTAGAAT	1356	
Db	1371	CGGGGCCCTGCTGGGGATCTCAAGGATGAGGCCCGCCAGCATGGAGATTGGGGGTAGAAT	1430	
QY	1357	TCCACCTTGAACCCAGACAGAGTCCTCCCTGCTGCTCCCTGACTCCCACTTGTAGGGTCTGC	1416	
Db	1431	TCCACCC-TGAACCCAGACAGAGTCCTCCCTGACTCCCACTTGTAGGGTCTGC	1489	
QY	1417	CTCAGGAGCCATAGAGTCCGGCTGTGT	1444	
Db	1490	CTCAGGAGCCATAGAGTCCGGCTGTGT	1517	
RESULT 11				
BC063553				
LOCUS				
DEFINITION Homo sapiens cDNA clone IMAGE:4939236, partial cds.				
ACCESSION BC063553				
VERSION BC063553.1 GI:39794649				
KEYWORDS				
SOURCE Homo sapiens (human)				
ORGANISM Homo sapiens				

REFERENCE	AUTHORS		
1 (bases 1 to 2034)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
1	Srausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Teshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Arambon, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wokley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Vallal, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 2034)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-DEC-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a> Tissue Procurement: David N. Louis, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc_mgc@hri.nih.gov">nisc_mgc@hri.nih.gov</a> Akhter, N., Avelle, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.D., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantipop, S., Thomas, P.J., Touchman, J.W., Tsurgion, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.		
source	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAC Plate: 134 Row: 0 Column: 3 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4885534. Location/Qualifiers 1. .2034 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4939236" /tissue_type="Brain, anaplastic oligodendroglioma with 1p/19q loss" /clone_lib="NCI CGAP_Brn67" /lab_host="DH10B" /note="Vector: PCMV-SPORT6"		
ORIGIN	Query Match 44.7%; Score 1204.4; DB 9; Length 2034; Best Local Similarity 99.5%; Pred. No. 96-280; 6; Indels 0; Gaps 0; Matches 1208; Conservative 0; Mismatches 6; Indels 0; Gaps 0;		



QY 277 CTTCTCTGTGAGCCAGCCCAAGATTTCAGGCGAGATGCCAGAGCACCCTGTCGTCCTCCACT 336  
Db  
QY 359 CTTCTCTGTGAGCCAGCCCAAGATTTCAGGCGAGATGCCAGAGCACCCTGTCGTCCTCCACT 418  
Db  
QY 337 GCGTAACTCTGCTGGTTCGAGAGACTGCCCGAAGGGAGGAGGACACACAGCCACGG 396  
Db 419 GCGTAACTCTGCTGGTTCGAGAGACTGCCCGAAGGGAGGAGGAGGACACACAGCCACGG 478  
QY 397 TGTAAAAACAGCCAGTGTGCTGTTCAATGGGACCCACAGGACCTGTGAGATCTGGAG 456  
Db 479 TGTAAAAACAGCCAGTGTGCTGTTCAATGGGACCCACAGGACCTGTGAGATCTGGAG 538  
QY 457 TTGGTCCAGTGGAGAGTGGCTGTGCTCGAGGCCCTGCTGGCCAGGCCAGAA 516  
Db 539 TTGGTCCAGTGGAGAGTGGCTGTGCTCGAGGCCCTGCTGGCCAGGCCAGAA 598  
QY 517 CTTCACTGTTTCATCAAAAACACAGTCACTTCAGCAAGTTCAACTTCTTAAGTCAA 576  
Db 599 CTTCACTGTTTCATCAAAAACACAGTCACTTCAGCAAGTTCAACTTCTTAAGTCAA 658  
QY 577 TGCCTTGGAGACTGGAGCCCACTATTTAAGCACTGCCCTATGAACACCAATTCAG 636  
Db 659 TGCCTTGGAGACTGGAGCCCACTATTTAAGCACTGCCCTATGAACACCAATTCAG 718  
QY 637 CCCCTACTGTCGGTGTTCGCAATTGGGACCTCTGTGGCAAGGCTGAGGAGACCTTCGA 696  
Db 719 CCCCTACTGTCGGTGTTCGCAATTGGGACCTCTGTGGCAAGGCTGAGGAGACCTTCGA 778  
QY 697 GGACCTGGCGTTGCTGGGTGGCTGTAGGCATCAGAGTTCACTGGGATTTGACCTGCA 756  
Db 779 GGACCTGGCGTTGCTGGGTGGCTGTAGGCATCAGAGTTCACTGGGATTTGACCTGCA 838  
QY 757 CACCGGGAGACTGGCTGTGGCTTCATCTCCCTTCAGTTCAGGAGAGAGCTACAA 816  
Db 839 CACCGGGAGACTGGCTGTGGCTTCATCTCCCTTCAGTTCAGGAGAGAGCTACAA 898  
QY 817 CTTCAAGGACAGCACTCACTGTGGAGCAACCGGTTGGAGGCCCGCCAGCTCTCAA 876  
Db 899 CTTCAAGGACAGCACTCACTGTGGAGCAACCGGTTGGAGGCCCGCCAGCTCTCAA 958  
QY 877 GCTCTATGGAATCCGCTTCGACATCTCTGTCACCGGGGAGGAGGAAAGTTTCGGGCTCAT 936  
Db 959 GCTCTATGGAATCCGCTTCGACATCTCTGTCACCGGGGAGGAGGAAAGTTTCGGGCTCAT 1018  
QY 937 CCCACGCGCTCAGCTGGGACACCGGGGAGCTTGGCTGGGCTGGTCACTTTTCTG 996  
Db 1019 CCCACGCGCTCAGCTGGGACACCGGGGAGCTTGGCTGGGCTGGTCACTTTTCTG 1078  
QY 997 TGACCTGCTACTGCTGTATGTGATAGAGAGCCCAATTTCTACTGGAGGACAAAGTATGA 1056  
Db 1079 TGACCTGCTACTGCTGTATGTGATAGAGAGCCCAATTTCTACTGGAGGACAAAGTATGA 1138  
QY 1057 GGAGGCCAAGGCCCGGAAAGCAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1116  
Db 1139 GGAGGCCAAGGCCCGGAAAGCAACCGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1198  
QY 1117 CCAAGCCGAGTGGCGGAGTGCTCAGACGGAGCTCAGCTCAGCCAGCCAGCCAGCTGC 1176  
Db 1199 CCAAGCCGAGTGGCGGAGTGCTCAGACGGAGCTCAGCTCAGCCAGCCAGCCAGCTGC 1258  
QY 1177 TGCTGGAGTCAAGACACAGACACCAAGGATGGCCCTGTCCAAAGTTCTGACACCCACTTGC 1236  
Db 1259 TGCTGGAGTCAAGACACAGACACCAAGGATGGCCCTGTCCAAAGTTCTGACACCCACTTGC 1318  
QY 1237 AACCCATTCGGGAGCTGTAGCG- TTCCCTGCTGTTTGA- -----GTGGGG 1286  
Db 1319 AACCCATTCGGGAGCTGTAGCGCTTCCCTGCTGTTGAGAGAGAGAGGGGCTGGGC 1378  
QY 1287 CTGGAGAGGGGGGGCCCTGCTGGGATCTCAAGGATGAGCCGCCAGCTATGA 1340  
Db 1379 AAGGAAGGAGCCCTGCCCTGCCGAGCAAGGATGAGCCAGCAATGA 1432

RESULT 13  
AR161587  
LOCUS AR161587 1293 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 1 from patent US 6255472.  
ACCESSION AR161587  
VERSION AR161587.1 GI:16227533  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1293)  
AUTHORS Tokino,T. and Nakamura,Y.  
TITLE Isolated nucleic acid molecule encoding a human skeletal  
muscle-specific receptor  
JOURNAL Patent: US 6255472-A 1 03-JUL-2001;  
FEATURES Location/Qualifiers  
source  
1..1293  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 42.9%; Score 1155.8; DB 6; Length 1293;  
Best Local Similarity 99.8%; Pred No. 4.9e-268;  
Matches 1157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 97 GTGGGCTCTCTCGCCAAAAAAGGCTACAGGAGCGGAGCTCGAAACCCAGTTTCCAT 156  
Db 135 GTGGGCTCTCTCGCCAAAAAAGGCTACAGGAGCGGAGCTCGAAACCCAGTTTCCAT 194  
QY 157 CATCACCAACTCAAGGGGTTTCCTCACTCAGATCAAGGACTTGGAAACCGGTGTG 216  
Db 195 CATCACCAACTCAAGGGGTTTCCTCACTCAGATCAAGGACTTGGAAACCGGTGTG 254  
QY 217 GGATGTGGCGCACTTCGTGAAGCCACTCAGGAGAGAAAGCTTCTTCTTGTGTGACCA 276  
Db 255 GGATGTGGCGCACTTCGTGAAGCCACTCAGGAGAGAAAGCTTCTTCTTGTGTGACCA 314  
QY 277 CTTCTTGTGACGCGCCCAAGTTTCAAGGAGATGCCAGAGACCCGCTCGTCCACT 336  
Db 315 CTTCTTGTGACGCGCCCAAGTTTCAAGGAGATGCCAGAGACCCGCTCGTCCACT 374  
QY 337 GGCTAACTGTGTCGAGGAGCTGCCCGAAGGGAGGAGGACACACAGCCACCG 396  
Db 375 GGCTAACTGTGTCGAGGAGCTGCCCGAAGGGAGGAGGAGGACACACAGCCACCG 434  
QY 397 TGTAAAAACAGGCGAGTGTGTGTTCATATGGACCCACAGGACCTGTGAGATCTGGAG 456  
Db 435 TGTAAAAACAGGCGAGTGTGTGTTCATATGGACCCACAGGACCTGTGAGATCTGGAG 494  
QY 457 TTGGTCCAGTGGAGAGTGGCGTTGTGCTGCTGAGGCCCTGCTGGCCAGGCCAGAA 516  
Db 495 TTGGTCCAGTGGAGAGTGGCGTTGTGCTGCTGAGGCCCTGCTGGCCAGGCCAGAA 554  
QY 517 CTTCACTGCTTTCATCAAAAAACACAGTCACTTTCAGCAAGTTTCAACTTCTTAAGTCCAA 576  
Db 555 CTTCACTGCTTTCATCAAAAAACACAGTCACTTTCAGCAAGTTTCAACTTCTTAAGTCCAA 614  
QY 577 TGCTTTGGAGACCTGGGACCCCACTTATTTAAGCATCTGGCTATGACCAAAATTCAG 636  
Db 615 TGCTTTGGAGACCTGGGACCCCACTTATTTAAGCATCTGGCTATGACCAAAATTCAG 674  
QY 637 CCCCTACTGTCCTGCTTCGGCATTTGGGACCTCTGTCGCAAGGCTGGAGGGACCTTCGA 696  
Db 675 CCCCTACTGTCCTGCTTCGGCATTTGGGACCTCTGTCGCAAGGCTGGAGGGACCTTCGA 734  
QY 697 GGAACCTGGGCTGCTGGTGGCTCTGTAGGATCAGAGTTCTCTGGATTTGACCTGGA 756  
Db 735 GGAACCTGGGCTGCTGGTGGCTCTGTAGGATCAGAGTTCTCTGGATTTGACCTGGA 794  
QY 757 CACCGGGAGCTGTGGCTGTGGCTCAGTACTCTTCCAGCTGCAGGAGAGAGCTACAA 816  
Db 795 CACCGGGAGCTGTGGCTGTGGCTCAGTACTCTTCCAGCTGCAGGAGAGAGCTACAA 854

QY 817 CTTACGACAGCCACTACTCTGTTGGAGCAACCGGGTGTGGAGCCCGCACCCCTGCTCAA 876  
Db 855 CTTACGACAGCCACTACTCTGTTGGAGCAACCGGGTGTGGAGCCCGCACCCCTGCTCAA 914  
QY 877 GCTCTATGGAATCCGCTTCGATCCTCGTCAACCGGGCAGCGAGGAGTTCGGGCTCAT 936  
Db 915 GCTCTATGGAATCCGCTTCGATCCTCGTCAACCGGGCAGCGAGGAGTTCGGGCTCAT 974  
QY 937 CCCACGCGCTGACACTGGGCAACCGGGCAGCTTGGCTGGGCTGTGCTCACTTTTCTG 996  
Db 975 CCCACGCGCTGACACTGGGCAACCGGGCAGCTTGGCTGGGCTGTGCTCACTTTTCTG 1034  
QY 997 TGACCTGCTACTGCTGTATGTTGGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1056  
Db 1035 TGACCTGCTACTGCTGTATGTTGGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1094  
QY 1057 GGAGGCCAAGCCCGGAAAGCAACCGGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1116  
Db 1095 GGAGGCCAAGCCCGGAAAGCAACCGGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1154  
QY 1117 CCAAGCCGACTGGCCGAGTGCCTCAGACGAGGCTCAGACCTGCACCCAGCGCACTGC 1176  
Db 1155 CCAAGCCGACTGGCCGAGTGCCTCAGACGAGGCTCAGACCTGCACCCAGCGCACTGC 1214  
QY 1177 TGCTGGGAGTCAGACACAGACAGGAGTGGCCCTGTCCAAAGTTCTGACACCCCACTTGC 1236  
Db 1215 TGCTGGGAGTCAGACACAGACAGGAGTGGCCCTGTCCAAAGTTCTGACACCCCACTTGC 1274  
QY 1237 AACCCATTCCGGAGCCTG 1255  
Db 1275 AACCCATTCCGGAGCCTG 1293

RESULT 14  
E17365  
LOCUS Human mRNA for P2XM protein. 1293 bp DNA linear PAT 28-JUL-1999  
DEFINITION  
E17365  
ACCESSION E17365  
VERSION 1 GI:5712048  
KEYWORDS JP 1998262681-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1293)  
Tokino, T. and Nakamura, Y.  
HUMAN GENE  
Patent: JP 1998262681-A 1 06-OCT-1998;  
OTSUKA PHARMACEUT CO LTD  
OS Homo sapiens (human)  
PN JP 1998262681-A/1  
PD 06-OCT-1998  
PP 26-MAR-1997 JP 1997093044  
PI TOKINO TAKASHI, NAKAMURA YUSUKE  
PC C12N15/09, A61K48/00, C07H21/04, C07K14/82//C07K16/32, C12P21/02,  
PC C12Q1/68,  
PC  
GOIN33/53, (C12N15/09, C12R1:91), (C12P21/02, C12R1:91), (C12P21/02, PC  
C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
FH key Location/Qualifiers  
FT source 1..1293 /organism='Homo sapiens'  
FT mat\_peptide 1..1293 /product='P2XM protein'.  
FT Location/Qualifiers  
1..1293  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

## ORIGIN

Query Match 42.9%; Score 1155.8; DB 6; Length 1293;  
Best Local Similarity 99.8%; Pred. No. 4.9e-268;  
Matches 1157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 97 GTGGGCTCTCCTCGCAAAAAGGCTACAGAGGCGGACCTGGAAACCCAGTTTCCAT 156  
Db 135 GTGGGCGCTCTCGCAAAAAGGCTACAGAGGCGGACCTGGAAACCCAGTTTCCAT 194  
QY 157 CATCACCAAACTCAAAGGGGTTTCCGTCACTCAGATCAAGGAGCTTGGAAAACCGCTGTG 216  
Db 195 CATCACCAAACTCAAAGGGGTTTCCGTCACTCAGATCAAGGAGCTTGGAAAACCGCTGTG 254  
QY 217 GGATGTGGCGCACTTCGTGAAGCCACCTCAGGAGAGAACGTTCTTCTTGTGTGACCAA 276  
Db 255 GGATGTGGCGCACTTCGTGAAGCCACCTCAGGAGAGAACGTTCTTCTTGTGTGACCAA 314  
QY 277 CTTCTTGTGAGCCAGCCCAAGTTTCAGGCGAGATGCCAGAGCACCCGTCGTCCTCACT 336  
Db 315 CTTCTTGTGAGCCAGCCCAAGTTTCAGGCGAGATGCCAGAGCACCCGTCGTCCTCACT 374  
QY 337 GGCTAACTGCTGGGTGCAAGGACTGCCCGAAGGGAGGAGGACACACAGCCACGG 396  
Db 375 GGCTAACTGCTGGGTGCAAGGACTGCCCGAAGGGAGGAGGACACACAGCCACGG 434  
QY 397 TGTAAAAACAGGCCCAAGTGTGTGGTGTTCATATGGGACCCACAGGACCTGTGAGATCTGGAG 456  
Db 435 TGTAAAAACAGGCCCAAGTGTGTGGTGTTCATATGGGACCCACAGGACCTGTGAGATCTGGAG 494  
QY 457 TTGGTCCCAAGTGGAGAGTGGCGTTGTGCCCTCGAGGCCCTGTGCGCCAGGCCAGAA 516  
Db 495 TTGGTCCCAAGTGGAGAGTGGCGTTGTGCCCTCGAGGCCCTGTGCGCCAGGCCAGAA 554  
QY 517 CTTACACCTGTTTCATCAAAAACACAGTCACCTTCAGCAAGTTCACTTCTCTAAGTCCAA 576  
Db 555 CTTACACCTGTTTCATCAAAAACACAGTCACCTTCAGCAAGTTCACTTCTCTAAGTCCAA 614  
QY 577 TGCCCTTGGAGACTCGGGACCCCACTATTTAAGCACTGCGCGCTATGAACACCAATTCA 636  
Db 615 TGCCCTTGGAGACTCGGGACCCCACTATTTAAGCACTGCGCGCTATGAACACCAATTCA 674  
QY 637 CCCCTACTGTCCCGTTCGCGATTTGGGGAACCTGTGTGCGCAAGGCTGAGAGGACCTTGA 696  
Db 675 CCCCTACTGTCCCGTTCGCGATTTGGGGAACCTGTGTGCGCAAGGCTGAGAGGACCTTGA 734  
QY 697 GGACCTGGCGTTGCTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTGTGACCTGGA 756  
Db 735 GGACCTGGCGTTGCTGGGTGGCTCTGTAGGATCAGAGTTCACTGGGATTGTGACCTGGA 794  
QY 757 CACCGGGGACTCTGCGCTGCTGCGCTCACTAATCTTTCAGCTGCAAGGAGAGAGTACAA 816  
Db 795 CACCGGGGACTCTGCGCTGCTGCGCTCACTAATCTTTCAGCTGCAAGGAGAGAGTACAA 854  
QY 817 CTTACGACAGCCACTCACTGTTGGAGCAACCGGGTGTGGAGGCCCGCACCCCTGCTCAA 876  
Db 855 CTTACGACAGCCACTCACTGTTGGAGCAACCGGGTGTGGAGGCCCGCACCCCTGCTCAA 914  
QY 877 GCTCTATGGAATCCGCTTCGATCCTCGTCAACCGGGCAGCGAGGAGTTCGGGCTCAT 936  
Db 915 GCTCTATGGAATCCGCTTCGATCCTCGTCAACCGGGCAGCGAGGAGTTCGGGCTCAT 974  
QY 937 CCCACGCGCTGACACTGGGCAACCGGGCAGCTTGGCTGGGCTGTGCTCACTTTTCTG 996  
Db 975 CCCACGCGCTGACACTGGGCAACCGGGCAGCTTGGCTGGGCTGTGCTCACTTTTCTG 1034  
QY 997 TGACCTGCTACTGCTGTATGTTGGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1056  
Db 1035 TGACCTGCTACTGCTGTATGTTGGATAGAGAGCCCATTTCTACTGGAGGACAAAGTATGA 1094  
QY 1057 GGAGGCCAAGCCCGGAAAGCAACCGGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1116  
Db 1095 GGAGGCCAAGCCCGGAAAGCAACCGGCAACTCTGTGTGGAGGAGCTGGCCCTTGCATC 1154



QY	1117	CGAAGCCCGAGTGC	CGCAGTGCCTCAGACGGAGCTCAGACCTGCACCCAGGCCACTGC	1117
Db	1155	CGAAGCCCGAGTGC	CGCAGTGCCTCAGACGGAGCTCAGACCTGCACCCAGGCCACTGC	1214
QY	1177	TGCTGGGAGT	CAGACACAGACACGAGTGGCCCTGTC	1236
Db	1215	TGCTGGGAGT	CAGACACAGACACGAGTGGCCCTGTC	1274
QY	1237	AACCCATTCGGGAGCCTG	1255	
Db	1275	AACCCATTCGGGAGCCTG	1293	
RESULT 15				
ABO10893				
LOCUS	ABO10893	1484 bp	mRNA	linear
DEFINITION	Mus musculus mRNA for mp2XM receptor, complete cds.			ROD 06-FEB-1998
ACCESSION	ABO10893			
VERSION	ABO10893.1	GI:2842547		
KEYWORDS	mp2XM; mp2XM receptor.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1484)			
AUTHORS	Nawa, G., Miyoshi, Y. and Nakamura, Y.			
TITLE	Cloning of mouse p2XM receptor gene, regulated by p53, and identification of its genomic structure			
JOURNAL	Published Only in Database (1998)			
REFERENCE	2 (bases 1 to 1484)			
AUTHORS	Nawa, G., Miyoshi, Y. and Nakamura, Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-FEB-1998) Yusuke Nakamura, Osaka University Medical School, Division of Clinical Genetics, Biomedical Research Center; 2-2, Yamadaoka, Suita, Osaka 565, Japan			
	(E-mail: REIG-NW@asahi-net.or.jp, Tel: 81-6-879-3381, Fax: 81-6-879-3389)			
FEATURES	Location/Qualifiers			
source	1. .1484			
gene	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/db_xref="taxon:10090"			
CDS	1. .1484			
	/gene="mp2XM"			
	45. .1184			
	/gene="mp2XM"			
	/codon_start=1			
	/product="mp2XM receptor"			
	/protein_id="BAA24693.1"			
	/db_xref="GI:2842548"			
	/translation="WASAAAAALVVGWGLDYKTVKLVNCRVGVSRILQLAVVYVIGWALLAKGQCELDLAPQTSVITKLKGVSVTOVKLENRLWDVDFPKPSQGENVFLTNLTVPAQVCGSPHPSVPLANCAWDEDCPEGTGTSGHKTGCVFNHTRTCEIWSVCPESGAVPRKPLAQAQNFLEIKNTVTFKSFNFSRNALLTWNTYFKHCLYDPLSPYCPFRIGDPLVAMAGDPELDLILGGAVGTLHWCNLTGKSGDCGPOYSQLOQKGNFRNTAHNWAASGVETRSLLLKGIREFDILVTGQAGKALIPPAITVGTGAHWLGWYFJCDLILLVLDREAGFYWKTYEARAPKTTTNS"			
ORIGIN				
Query Match	28.9%	Score 777;	DB 10;	Length 1484;
Best Local Similarity	76.6%	Pred. No. 1.6e-176;		
Matches 994;	Conservative 0;	Mismatches 290;	Indels 13;	Gaps 3;
QY	97	GTGGGCTCTCTCGCCAAAAAGGCTTACCAGAGCGGACCTGGAAACCCAGATTTTCCAT	156	
Ddb	185	GTGGGCTCTCTTGTGCCAAAAGAGCTACAGGAACGGGACTTGGCCCTCAGACTTCTGT	244	
QY	157	CATCACAAACTCAAGGGGTTTCCTGCTACTCAGATCAAGAGCTTGGAAACCGGTGTG	216	
Ddb	245	CATCACAAACTCAAGGGGTTTCTGTAAACCCAGGTTAAGAGCTGGAGAACCGGTGTG	304	
QY	217	GGATGTGGCCGACTTCGTGAAGCCACTCAGGAGAGAACTGTCTTCTTGTGTGACCAA	276	

Db 1384 TTGGGAAGGAAGAGGGTTCTGGCTTGGGAAATGGAGGACAAAGCTTCAGCGGACTAGA 1443  
Qy 1345 TGGGGGTAGAAATCCACCCCTTGAACCCCGACGACAG 1381  
Db 1444 GGGGAGTTGGGTAATTCGACAGACTCCGATGCAG 1490

Search completed: November 22, 2004, 01:14:55  
Job time : 11532.5 secs